Journal of Consulting and Clinical Psychology

Therapeutic Bond Judgments: Congruence and Incongruence

Dana Atzil-Slonim, Eran Bar-Kalifa, Eshkol Rafaeli, Wolfgang Lutz, Julian Rubel, Ann-Kathrin Schiefele, and Tuvia Peri

Therapeutic Bond Judgments: Congruence and Incongruence

Dana Atzil-Slonim, Eran Bar-Kalifa,
and Eshkol Rafaeli
Bar-Ilan University

Wolfgang Lutz, Julian Rubel, and
Ann-Kathrin Schiefele
University of Trier

Tuvia Peri
Bar-Ilan University

Objective: The present study had 2 aims: (a) to implement West and Kenny's (2011) Truth-and-Bias model to simultaneously assess the temporal congruence and directional discrepancy between clients' and therapists' ratings of the bond facet of the therapeutic alliance, as they cofluctuate from session to session; and (b) to examine whether symptom severity and a personality disorder (PD) diagnosis moderate congruence and/or discrepancy. ***Method:*** Participants included 213 clients treated by 49 therapists. At pretreatment, clients were assessed for a PD diagnosis and completed symptom measures. Symptom severity was also assessed at the beginning of each session, using client self-reports. Both clients and therapists rated the therapeutic bond at the end of each session. ***Results:*** Therapists and clients exhibited substantial temporal congruence in their session-by-session bond ratings, but therapists' ratings tended to be lower than their clients' across sessions. Additionally, therapeutic dyads whose session-by-session ratings were more congruent also tended to have a larger directional discrepancy (clients' ratings being higher). Pretreatment symptom severity and PD diagnosis did not moderate either temporal congruence or discrepancy at the dyad level; however, during sessions when clients were more symptomatic, therapist and client ratings were both farther apart and tracked each other less closely. ***Conclusions:*** Our findings are consistent with a "better safe than sorry" pattern, which suggests that therapists are motivated to take a vigilant approach that may lead both to underestimation and to attunement to fluctuations in the therapeutic bond.

What is the public health significance of this article?
This study advances the idea that therapists who adopt a vigilant approach may be more attuned to their clients' changing experience, Additionally, it highlights the risk of misattributing symptomatic change to factors within the therapeutic relationship.

Clients' perceptions of the therapeutic alliance matter (Horvath, Del Re, Flückiger, & Symonds, 2011), but so do therapist perceptions (Del Re, Flückiger, Horvath, Symonds, & Wampold, 2012). Yet, these perceptions may be only partly congruent (Tryon, Blackwell, & Hammel, 2007). In fact, an interesting paradox emerges regarding the congruence between clients' and therapists' judgments of alliance: When therapists are asked to rate the therapeutic alliance, they provide ratings that are moderately and positively correlated with their clients'

ratings (cf., Marmarosh & Kivlighan, 2012; Tryon et al., 2007); at the same time, therapists, compared with clients, tend to be biased toward underestimating the alliance (Fitzpatrick, Iwakabe, & Stalikas, 2005; Tryon et al., 2007).

Studies of alliance congruence have typically reported data relevant to either correlations or mean-level differences between clients' and therapists' ratings. These studies have rarely utilized successive repeated measures from both parties in the therapeutic relationship to assess congruence at the within-dyad level—a type of congruence that we will refer to as temporal congruence. Even studies that have collected repeated measures (Fitzpatrick et al., 2005; Hilsenroth, Peters, & Ackerman, 2004; Kivlighan & Shaughnessy, 1995) have tended to use a limited number of sessions (typically 2 or 3), reflecting phases within therapy rather than changes from session to session. Consequently, existing knowledge addresses the broad question of congruence between clients and therapists across dyads and not the issue of congruence across sessions *within* the same dyads.

Dana Atzil-Slonim, Eran Bar-Kalifa, and Eshkol Rafaeli, Department of Psychology, Bar-Ilan University; Wolfgang Lutz, Julian Rubel, and Ann-Kathrin Schiefele, Department of Psychology, University of Trier; Tuvia Peri, Department of Psychology, Bar-Ilan University.

Dana Atzil-Slonim and Eran Bar-Kalifa equally contributed to this article.

Correspondence concerning this article should be addressed to Dana Atzil-Slonim, 20 HaBanay st. Jerusalem, Israel. E-mail: dana.slonim@gmail.com

The current investigation is the first within the study of therapeutic alliance to apply an innovation in dyadic judgment research: West and Kenny's (2011) recently developed Truth-and-Bias (T&B) model. Using this model, we will (a) assess, session-by-session across the initial months of treatment, the temporal congruence and directional discrepancy between the client's and the therapist's experience of their bond; (b) examine the association between the congruence and discrepancy measures; and (c) examine whether clients' diagnoses and symptom levels are moderators of congruence and/or discrepancy.

It is important to clarify the definition of alliance we use in this study, as this concept has become increasingly complex over time. Bordin's (1994) tri-partite model of alliance views it as involving goal consensus, task consensus, and the interpersonal bond. Of these, the bond is often seen as the core of the alliance (e.g., Hartmann, Joos, Orlinsky, & Zeeck, 2014). As previous reviews of the alliance literature (e.g., Horvath et al., 2011; Orlinsky & Howard, 1986) have noted, the *bond* refers specifically to the relationship between client and therapist as persons, and is distinct from the "therapeutic contract" that organizes their distinct roles (and that corresponds more closely to Bordin's *goal* and *task* concepts). In our work, we focus on the congruence between therapists' and clients' ratings of the therapeutic bond rather than other facets of the therapeutic alliance.

Congruence and Discrepancy in Client-Therapist Alliance Ratings

Congruence between therapists' and clients' views of the therapeutic alliance is considered important both when alliance is strong and when it is weak (Horvath et al., 2011). When clients and therapists agree that the alliance is strong, they work in concert to address the client's needs. When alliance rupture occur, the therapist's rapid recognition of the occurrence is important because it may allow the therapist to take appropriate action (Safran & Muran, 1996).

Most studies on client-therapist congruence have used single-session measures of alliance (Marmarosh & Kivlighan, 2012), collapsed alliance ratings across sessions to form "phases" of treatment (Fitzpatrick et al., 2005; Hilsenroth et al., 2004), or selected several sessions, typically from various phases, as representative (Kivlighan & Shaughnessy, 1995). In most studies, congruence has been conceptualized as the correlation between therapists' and clients' alliance ratings or as the mean differences between these ratings.

Studies that define congruence as the correlation between clients' and therapists' ratings of the alliance (e.g., Kivlighan & Shaughnessy, 1995) have generally found a moderate association between these ratings. In a meta-analysis of these studies, Tryon et al. (2007) reported that in 32 studies with correlational congruence data, the mean correlation was $r = .36$, with little variability across studies. What should we think of a correlation of this magnitude? It may suggest that only a small amount ($\sim$13%) of the variability in clients' alliance ratings can be predicted by their therapists' assessments of the alliance. However, research on interpersonal perception in various areas (cf., Kenny, 2004) should comfort us somewhat. As Kenny and others have shown, even well-acquainted perceivers are typically congruent only at moderate or even lower levels.

Studies that define congruence as the absolute or relative difference between clients' and therapists' ratings of the alliance (e.g., Fitzpatrick et al., 2005) have shown that, on average, clients' ratings were significantly higher than therapists' ratings. In their meta-analysis, Tryon et al. (2007) reported that in 44 studies reporting difference scores, the mean difference was $d = 0.63$ ($SD = 0.42$). Although many studies simply provide data attesting to this directional bias, a few authors have speculated about possible reasons for this consistent finding. A primary theory is that therapists have experienced alliance with other clients and, therefore, may rate their alliance with the focal client in relation to alliances with previous clients. In contrast, clients may have little prior experience in therapy and, therefore, less ability to anchor their judgments about their current alliances (Fenton, Cecero, Nich, Frankforter, & Carroll, 2001). Because of this novelty, clients may compare their therapeutic alliance to their experience with other health service professionals (e.g., medical doctors) who often adopt less collaborative approaches compared with psychotherapists (Tryon et al., 2007). Another suggestion (Bachelor & Salame, 2000; Fitzpatrick et al., 2005) is that research studies often utilize trainee therapists who tend to be more self-critical and less confident in their abilities, although their clients tend to perceive their work quite favorably.

A recent study that used more sophisticated methods to examine therapist–client discrepancy (Marmarosh & Kivlighan, 2012) raised the possibility that therapists' underestimation of the alliance can be adaptive for the relationship. In this study, clients rated sessions as less smooth when their ratings of the alliance were lower than their therapists' ratings of the alliance. The authors argued that when therapists overestimate alliance, they may fail to address ruptures and, therefore, may not take appropriate action to repair them. In contrast, if therapists underestimate alliance, fewer negative consequences ensue—essentially, better safe than sorry.

Remaining Questions Regarding Congruence in Alliance Ratings

Tryon et al. (2007) called for more studies addressing the factors that lead to therapist and client congruence or incongruence in alliance ratings. Notably, even recent and sophisticated studies on the topic (e.g., Marmarosh & Kivlighan, 2012) have not been based on session-by-session data. These data would allow us to address three questions that, in our view, have received insufficient attention in the extant alliance congruence literature.

First, this literature has yet to examine temporal congruence. Should we expect such congruence to be moderate on average, as has been found at the between-dyad level (cf., Tryon et al., 2007)? Moreover, will there be between-dyad variability in such congruence, and if so, what would moderate it? Second, although the extant literature suggests a mean negative discrepancy, this finding has been based on single-session or pooled alliance ratings. Should we expect the same therapist underestimation to occur over time? Moreover, will there be between-dyad variability in this discrepancy, and if so, what would moderate it? Session-by-session data allows us to examine both person-level moderators (e.g., the client's psychopathology) and session-level moderators (e.g., session-level symptomatology) for temporal congruence and for discrepancy. Third, the extant congruence literature has studied temporal congruence and mean-level discrepancy independently of

each other. To the best of our knowledge, no study has explored the association between the two; had these associations been computed, they would not have been based on indices of temporal tracking. Therefore, a lingering question is whether within-dyad temporal congruence decreases as bias increases or, rather, whether the same therapists are both incongruent (i.e., show negative discrepancy) and congruent (i.e., show temporal agreement) with their clients.

Lessons From Dyadic Research

The concepts of temporal congruence (tracking accuracy) and mean-level discrepancy (directional bias) in interpersonal judgments have received attention in other fields of psychology, especially among dyadic relationship researchers. A recent review (Fletcher & Kerr, 2010) showed that tracking accuracy and mean-level bias in the judgment of relationship characteristics (e.g., supportiveness, conflict) can co-occur. However, although 23 of the studies reviewed in Fletcher and Kerr's meta-analysis provided correlations between partners across multiple times or situations (and could have yielded mean-level directional differences between partners), none reported a link between tracking accuracy and directional bias, possibly because these constructs were always obtained in separate analyses.

In their recently developed T&B model, West and Kenny (2011) presented a statistical innovation that allows mean-level bias and tracking accuracy to be assessed simultaneously in a single analysis using the same measures and benchmarks. Therefore, researchers can estimate the association between the two. Specifically, accuracy and bias may be positively, negatively, or not at all related depending on which psychological factors influence judgment.

In one application of the T&B model, Overall, Fletcher, and Kenny (2012) assessed the association between tracking accuracy and bias in participants' judgments about their romantic partner's regard for them. They found that perceivers who underestimated their partner's regard also tracked it more accurately over time. They interpreted this finding as support for the idea that a vigilant (i.e., better safe than sorry) approach regarding one's relationship can underlie both negative bias and tracking accuracy. This interpretation is consistent with Haselton and Buss's (2000) error management theory, an evolutionary psychology model that suggests that in particular contexts, false-negative judgments (e.g., underestimation) are less costly than false-positive ones; judgments regarding the connection with another person may be one such context. Strong relationships should, in principle, involve the motivation to accurately track one's partner's thoughts and feelings. However, underestimating a partner's forgiveness, trust, or love is safer than overestimating it. Specifically, the latter might lead to complacency and a lack of effort in attempting to repair problems in the relationship.

Applying These Lessons to the Study of Bond Congruence

In our study, we examine whether the same "better safe than sorry" pattern found by Overall et al. (2012) and by Fletcher and Kerr (2010) may also occur with regards to congruence in ratings of the bond facet of the therapeutic alliance. We apply West and Kenny's (2011) T&B model to clients' and therapists' judgments of their bond. By assessing therapists' ratings repeatedly across treatment, with clients' bond ratings serving as a benchmark, we can simultaneously assess temporal congruence (i.e., whether a therapist's ratings tracked their client's changing bond ratings) and mean-level discrepancy (i.e., whether the therapist's ratings were positive or negatively biased vis-à-vis the client's). In T&B terms, temporal congruence can be considered a *truth force*, and mean-level discrepancy can be considered a *directional bias*.

Prototypical T&B studies use person A's ratings as the truth criterion, and instruct person B to infer or judge person A's ratings. The present study departs from this design, exploring congruence or agreement within the therapeutic dyad rather than accuracy or bias. Still, therapy involves an asymmetrical relationship in which one of the therapist's explicit goals is to attend to and understand their clients' experience, but not vice versa. Therefore, we treated the clients' ratings as the "truth" benchmark and the therapists' ratings as the (to-be-predicted) "judgment." This choice does not imply that the clients' view is any truer (or less true) than the therapists'.

More important, because the T&B framework is able to examine discrepancy (directional bias) and temporal congruence (tracking accuracy) simultaneously, it lets us examine the degree to which the two are related. Conceptually, the two may be independent of each other. As Figure 1 illustrates, particular therapeutic dyads (drawn from our sample) may demonstrate discrepancy and congruence that are high-high, low-high, high-low, or low-low. However, if the better safe than sorry pattern found in dyadic research is also valid within therapeutic dyads, we may expect dyads with high congruence to have high negative discrepancy, and vice versa.

Moderators of Congruence and Discrepancy

The T&B model allows us to include moderators of both temporal congruence and mean-level discrepancy. One possible moderator is the clients' level of distress. Indeed, of the several possible moderators explored by Tryon et al. (2007), client distress received the clearest support. When clients experience more severe symptoms, they and their therapists may be more congruent in estimating the bond or alliance for one of several reasons. For one, distress is likely to be tied to greater negativity, thus, bringing the client's view closer to the (perennially more pessimistic) therapist view. Alternatively, greater distress may serve as a stronger signal for therapists, who can then base their own estimates on these signals from their clients.

More important, the data meta-analyzed by Tryon and her colleagues were sample-level effect sizes. These allowed the authors to reach conclusions about sample-level moderators—for example, mean sample-level therapist/client discrepancy was found to be smaller in *samples* in which clients had severe symptoms than in ones in which they had either moderate or mild symptoms. In contrast, their data did not allow examining person-level or session-level moderation of congruence; that is, they cannot indicate whether mean-level congruence or discrepancy differ for *clients* with worse symptoms or with a personality disorder (PD) diagnosis, nor can they indicate the extent to which higher symptomatology moderates congruence or discrepancy at the session-level.
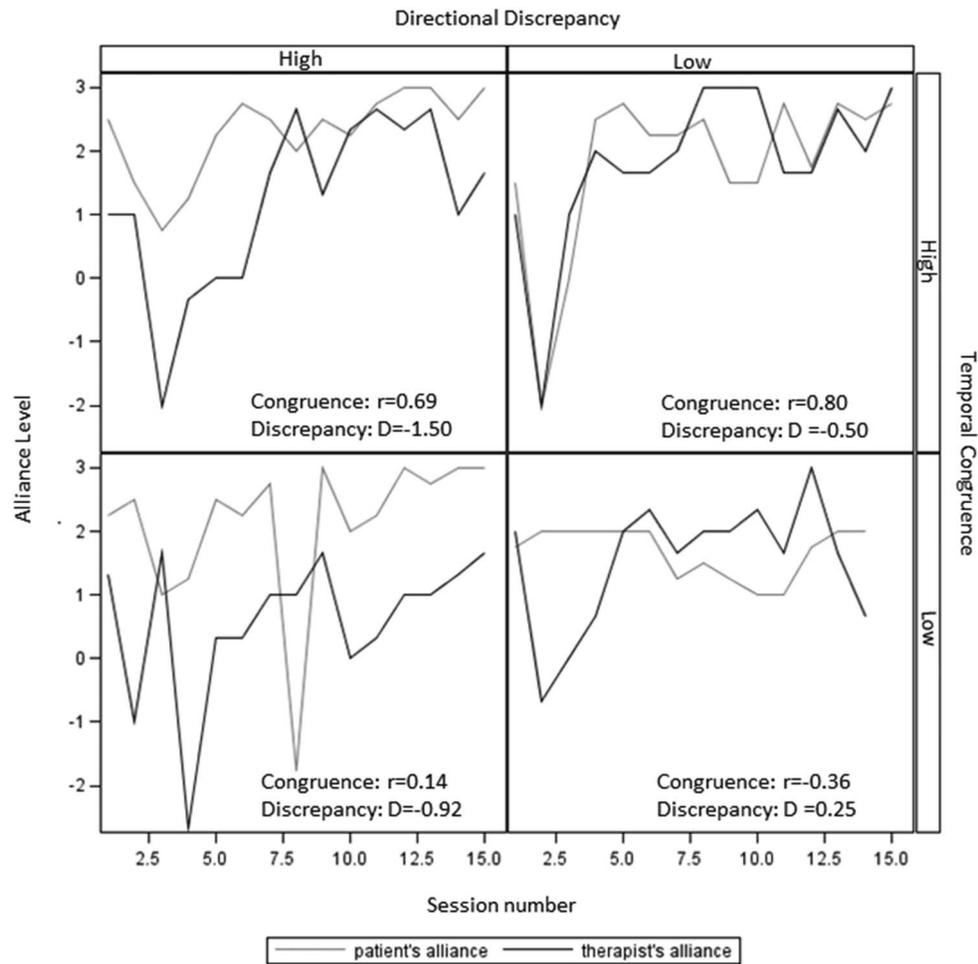
Figure 1. Illustration of the four possible patterns of discrepancy/congruence relations drawn from specific therapist-client dyads in our sample.

In some ways, the association between distress and congruence or discrepancy may be tied to the more basic associations between distress and the alliance ratings themselves. For example, distress may bring the client's alliance ratings down, and consequently, make them closer to the therapists' ratings. Of interest to the authors, the evidence for any association between pretreatment distress and either reporter's alliance ratings is mixed. For example, Hersoug, Høglend, Havik, and Monsen (2010) found that pretreatment global functioning was related to growth in the alliance over time for those with more adaptive functioning. In contrast, several authors found no significant association between pretreatment symptoms and alliance ratings (e.g., Connolly Gibbons et al., 2003). Studies have only recently begun to examine the association between alliance and symptoms at the session level, and evidence was found for a reciprocal causal model, in which client-rated alliance predicted subsequent change in symptoms and vice versa (Falkenström, Granström, & Holmqvist, 2013).

Mixed findings were also found in studies examining the association between alliance and a different proxy index of distress—namely, PD diagnosis. A number of clinical researchers have noted that because PD clients present with longstanding and inflexible patterns of emotional and interpersonal difficulties, they invariably pose great challenges to therapists, especially with regard to the therapeutic alliance (e.g., Muran et al., 2009). Other authors have shown that, at least in borderline personality disorder (BPD), moment-to-moment changes and shifts in the alliance are common (Levy, Beeney, Wasserman, & Clarkin, 2010). In contrast, several studies have found no significant association between PD diagnosis and alliance ratings (e.g., Hersoug et al., 2010; Tufekcioglu, Muran, Safran, & Winston, 2013).

Given these mixed findings regarding the association between client distress and alliance ratings, and the limited findings regarding distress and alliance congruence, we will explore indices of person-level distress (PD diagnosis and pretreatment symptoms) as well as session-level symptoms as possible moderators of congruence and discrepancy in clients' and therapists' ratings of the bond facet of the therapeutic alliance.

Research Questions and Hypotheses

This study was organized around two broad goals. As a first goal, we wished to examine whether therapists' judgments of the

therapeutic bond are congruent with those of the clients. Unlike previous research on congruence, we utilized West and Kenny's (2011) T&B model, which allowed us to determine whether therapists are discrepant in their levels and/or congruent over time. This goal comprises several hypotheses.

Prerequisite Hypothesis

We expect to find significant variability in clients' and therapists' ratings of the bond. Similar findings have been identified in several studies (e.g., Bachelor & Salame, 2000); replicating these is a prerequisite for the subsequent hypotheses.

Primary Hypothesis 1a

We expect that therapists' bond ratings will generally be lower than their clients' ratings, session by session. This prediction is consistent with findings from both psychotherapy research (Tryon et al., 2007) and relationship research (Fletcher & Kerr, 2010; Overall et al., 2012). The use of session-by-session bond ratings from the initial months of treatment will allow us to examine directional bias as a session-level phenomenon. Additionally, it will allow us to examine Hypotheses 1b and 1c, described below.

Primary Hypothesis 1b

We expect therapists' ratings to be temporally congruent with their clients' bond ratings, based on previous congruence studies (Tryon et al., 2007). Notably, congruence in these earlier studies was not assessed on a session-by-session basis.

Primary Hypothesis 1c

We expect therapists who are more negatively biased to also demonstrate greater temporal congruence. This prediction is based on error management theory (Haselton & Buss, 2000) and on findings from relationship research showing that perceivers who were more negatively biased tracked their partner's regard more accurately (Overall et al., 2012).

Hypotheses 2a–2f

For the study's second goal, we wished to examine whether client distress moderated discrepancy or temporal congruence between therapists' and clients' bond ratings. We examine this moderation both at the person and at the session levels. The *person-level* predictions were that PD diagnosis (Hypothesis 2a) and/or symptom severity (Hypothesis 2b) would moderate discrepancy, such that therapists of clients diagnosed with PD and/or showing higher pretreatment severity will show less (even if still negative) directional bias. These hypotheses are consistent with Tryon et al.'s (2007) finding of less directional bias in *samples* characterized by greater severity. Additionally, we expected that PD diagnosis will be associated with lower temporal congruence (Hypothesis 2c). Though no existing studies inform us about this association, we speculate that individuals with this diagnosis, who tend to fluctuate in their experience, also fluctuate in their bond ratings, making it more difficult for therapists to be congruent. The closest available evidence supporting this reasoning comes from Levy et al., (2010), who showed that mental state vacillations in

clients with BPD led to poorer therapeutic alliance. Finally (exploratory Hypothesis 2d), we had no basis for predicting a particular association between pretreatment symptom severity and temporal congruence; we explore it in our analyses with no directional prediction.

The *session-level* predictions had little previous research for reference. Therefore, both of the predictions (regarding discrepancy and temporal congruence) are exploratory. First, (exploratory Hypothesis 2e), we will test whether session-level symptoms moderate discrepancy, with sessions characterized by higher symptomatology having less (negative) directional bias. Such a pattern, which would be similar to the one predicted at the person level, could be expected based on the arguments presented by Tryon et al. (2007). Although no previous studies have examined congruence at the session-by-session level, studies examining the association between client-rated *alliance* and symptoms have found the two to be negatively associated (Falkenström et al., 2013). Specifically, if symptoms worsen, clients tend to rate the alliance lower. As we reasoned above, this process is likely to reduce the gap between the (perennially pessimistic) therapist and the (situationally glum) client. Notably, however, therapists may also become glum if they encounter a worsening of symptoms in particular sessions, and they may see these worsened symptoms as associated with a worsened therapeutic bond. If we were to apply the better safe than sorry rationale, we would actually expect therapists to react to these threatening situations by further underestimating the bond, which may cause the directional bias to remain intact or even increase. Given the limited evidence for either prediction, we treat this examination as exploratory. Similarly (exploratory Hypothesis 2f), as we have no basis for predicting a particular association between symptoms at the session level and temporal congruence, we explore these too with no directional prediction.

Method

Participants and Treatment

The analyses were based on a sample that comprised 213 clients treated by 49 therapists at the Trier University Outpatient Clinic in Southwest Germany between 2009 and 2013. All therapists in this project participated in a 3-year (full-time) or 5-year (part-time) postgraduate training program with a cognitive–behavioral therapy (CBT) focus and had at least 1 year of training before beginning to see clients. Clients included in the analysis had at least eight sessions of individual treatment with a mean treatment length of 39 sessions ($SD = 14.68$, range $= 8–87$). Data from the first 15 sessions (if available) were used for the current analyses ($M = 13.42$, $SD = 3.20$, range 8–15). Clients were over 15 years old ($M = 36.26$, $SD = 13.06$, range $= 15–74$), and the majority was female (66.7%). Diagnoses were based on the S*tructured Clinical Interview for Axis I DSM–IV Disorders—Patient Edition* (*SCID-I*; First, Spitzer, Gibbon, & Williams, 1995). Most clients were diagnosed with an affective disorder (53%) or anxiety disorder (20.2%) as the primary diagnosis. Additional primary diagnoses were acute stress and adjustment disorders (12.2%), obsessive–compulsive disorder (3.8%), eating disorders (2.4%), somatoform disorders (2.3%), and others (6.1%). For the diagnosis of

personality disorders, the *International Diagnostic Checklist for Personality Disorders* (*IDCL-P*; Bronisch, Hiller, Mombour, & Zaudig, 1996) was adopted, which identified 41 clients as having a personality disorder (19.2%). Both interviews were conducted before the actual therapy by intensively trained independent clinicians. These interviews were videotaped; interviews and diagnoses were discussed in expert consensus teams that included four senior clinicians; final diagnoses were determined by consensual agreement of at least 75% of the team members.

Instruments and Data Collection

Brief symptom inventory (BSI). To assess symptom severity, the BSI (Franke, 2000; German translation of Derogatis, 1975) was administered pretreatment. This 53-item self-report inventory inquires about physical and psychological symptoms within the last week. It is the brief form of the Derogatis' Symptom Check-List-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 psychoticism. Items are based on a 5-point Likert scale ranging from 0 (*not at all*) to 4 (*extremely*). In this study only the Global Severity Index (GSI) was used. The psychometric properties of this index can be regarded as excellent ($\alpha = .92$; $r_{tt} = .90$; see Franke, 2000).

Hopkins symptom checklist-short form (HSCL-11). The HSCL-11 (Lutz, Tholen, Schürch, & Berking, 2006) was administered at the beginning of each session. This 11-item self-report inventory assesses symptomatic distress. It is a brief version of the SCL-90-R (Derogatis, 1992). The items are based on a 4-point Likert scale ranging from 1 (*not at all*) to 4 (*extremely*). The mean of the 11 items represents the client's level of global symptomatic distress in the preceding week. It is highly correlated with the GSI ($r = .91$) and 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. Based on Grawe (1997) dual model of psychotherapy, and on Tschacher, Ramseyer, and Grawe's (2007) factor analysis, we chose to use only the *Global Alliance* subscale, which includes four items in the patient version and three items in the therapist version.[1] Items in both are answered on 7-point Likert scales ranging from −3 (*not at all*) to 3 (*yes, exactly*). Both versions had excellent internal consistency (for patients, α ranged between .87 to .92 for different sessions; for therapists, it ranged between .84 to .87). 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). The correlation between the BPSR *global alliance* subscale and the *relationship satisfaction* scale of the *Revised Helping Alliance Questionnaire* (HAq-II; Luborsky et al., 1996) for the whole sample of clients treated at the outpatient clinic from which the data are obtained were: $r = .58$ ($N = 845$) in Session 5, $r = .62$ ($N = 845$) in Session 10, and $r = .58$ ($N = 669$) in Session 15. These associations strengthen the evidence for the validity of the BPSR subscale as a therapeutic bond measure.

Data Analytic Strategy

The data set had a hierarchical structure in which session ratings were nested within clients, who were nested within therapists. As a result, individual observations were not independent of each other, and the session-by-session ratings of any client were more similar to each other than those of two randomly chosen clients. For this reason, hierarchical data violate the assumption of independent observations made by traditional statistical methods, but may be handled appropriately using multilevel modeling (MLM; Hox, 2010). To analyze the present data set, a three-level MLM was used, partitioning the total variability in bond ratings for session s of client c who was treated by therapist t into three components: variance within client or therapist ratings at Level 1, between clients at Level 2, and between therapists at Level 3.

Results

Variability in Clients' and Therapists' Bond Ratings

To test our prerequisite hypothesis that *clients'* bond ratings would show variability at the session level and that this variability would be more prominent than variability across clients or therapists, we conducted a three-level unconditional model. The three equations were, respectively,

Level 1:

$$\text{Judgment}_{sct} = \pi_{0ct} + e_{sct};$$

Level 2:

$$\pi_{0ct} = \beta_{00t} + r_{0ct};$$

Level 3:

$$\beta_{00t} = \gamma_{000} + u_{00t}$$

where the bond rating for session s of client c who was treated by therapist t was modeled at Level 1 as a function of the client mean bond rating (π_{0ct}) plus a residual term quantifying the specific-session deviation around this mean (e_{sct}). At Level 2, the client mean was modeled as a function of the mean bond ratings of all clients treated by therapist t (β_{00t}) plus a residual term quantifying the specific-client deviation around this mean (r_{0ct}). At Level 3, the mean bond ratings of the clients who were treated by therapist t were modeled as a function of the sample

---

[1] The BPSR Global Alliance subscale, patient version, comprises four items: "The relationship with my therapist felt comfortable today," "My therapist and I are getting along well," "I think my therapist is genuinely concerned about my wellbeing," "I feel that the therapist has real appreciation for me." The BPSR Global Alliance subscale, therapist version comprises three items: "The relationship with my patient felt comfortable today," "My patient and I are getting along well," "My patient and I are collaborating on the same goals." To examine whether discrepancies in ratings were because of the use of different items for patients and therapists, we also constructed briefer subscales for both the patient and therapist versions; these consisted of the first two items in each version, which are worded identically for both responders. Analyses that were ran with these shortened subscales yielded similar results to the analyses that were ran with the full subscales. Thus, in the current study we report the analyses conducted with the full subscales to allow comparisons across studies.

Table 1
Unconditional Models for Clients' and Therapists' Alliance Ratings

Parameter estimates	Unconditional model for clients' alliance			Unconditional model for therapists' judgments		
	Estimate (*SE*)	*p* value	Effect size[a]	Estimate (*SE*)	*p* value	Effect size[a]
Fixed effects						
Intercept (γ_{000})	2.23 (0.04)	<0.001**	0.99	1.69 (0.07)	<0.001**	0.94
Random effects						
Level 1 (sessions)						
Residual	0.43 (0.01)	<0.001**		0.57 (0.01)	<0.001**	
Level 2 (clients)						
Intercept	0.29 (0.03)	<0.001**		0.20 (0.03)	<0.001**	
Level 3 (therapists)						
Intercept	0.01 (0.02)	0.223		0.15 (0.05)	0.001*	
Model summary						
−2 LogL (deviance)	6717.5			7478.4		
No. estimated parameters	4			4		

Note. *p* values for fixed effects were based on two-tailed *t*-tests with Kenward & Roger (1997) correction method for computing DF; *p* values for random effects were based on one-tailed Wald z test because variances are constrained to be nonnegative.
[a] Effect sizes were estimated with semipartial R^2 for linear mixed models (Edwards et al., 2008).
* $p < .01$. ** $p < .001$.

grand mean (γ_{000}) plus a residual term quantifying the deviation of this therapist's clients from the grand mean (u_{00t}). The results of these equations are presented on the left side of Table 1. The sample mean bond was 2.23 (on a scale ranging from −3 to 3), indicating that clients tended to rate the therapeutic bond as very high. Additionally, as predicted, there was a significant variability at the session level (i.e., Level 1). There was also significant variability at the client level (i.e., Level 2) but no significant variability at the therapist level (i.e., Level 3). Using the variability estimates to calculate the proportion of variation at each level, we found that 59.9% of the variability was at the session level, 39.7% was at the client level, and 1.4% was at the therapist level.

To test our prerequisite hypothesis that *therapists'* bond ratings would also show variability at the session level and that this variability would be more prominent than the variability across clients or therapists, we conducted a similar three-level unconditional model with the therapists' ratings of the bond as the outcome. The results of these equations are presented on the right side of Table 1. The therapists sample mean ratings of bond were 1.69, indicating that therapists tended to rate the bond highly, although not as highly as their clients. As predicted, there was significant variability at the session level (62.0%). There was also significant variability at the client level (21.7%) and at the therapist level (16.3%).

Bond Ratings' Congruence, Discrepancy, and Their Association

To test our primary hypotheses regarding congruence and discrepancy in clients and therapists ratings of the bond, we used West and Kenny's (2011) T&B model. Here, therapists' reports of the bond constituted the *judgment* and served as the outcome. The judgment was predicted by the clients' reports of the bond, which constituted the *truth*. The slope coefficient of the truth (i.e., the *truth force*) represents the degree to which clients' and therapists'

ratings were congruent. As West and Kenny (2011, pp. 374–375) suggest, we centered the judgment and truth parameters around each person's mean truth score (i.e., the client's mean bond ratings) across all sessions. This allowed us to remove broad individual differences when examining within-person fluctuations. It also allowed the intercept estimate to represent the *directional bias* (i.e., the degree to which therapists overestimate (in cases of positive intercepts) or underestimate (in cases of negative ones) the bond compared with their clients' experience). We examined a three-level model in which directional bias (i.e., the intercept) was treated as random at both Level 2 (i.e., the client level) and Level 3 (i.e., the therapist level), and the truth variable (i.e., clients' ratings) was treated as random at Level 2.[2] The model's equations were

Level 1:

$$\text{Judgment}_{sct} = \pi_{0ct} + \pi_{1ct} * \text{Truth}_{sct} + e_{sct}$$

Level 2:

$$\pi_{0ct} = \beta_{00t} + r_{0ct}; \pi_{1ct} = \beta_{10t} + r_{1ct}$$

Level 3:

$$\beta_{00t} = \gamma_{000} + u_{00t}; \beta_{10t} = \gamma_{100}.$$

Table 2 displays the results of the T&B analysis including fixed and random effects. The fixed directional bias was negative and significant, indicating that, as predicted (Hypothesis 1a), therapists tended to rate the bond lower than their clients. The random effect of the directional bias was also significant at Level 2, indicating significant between-client variability in the extent to which therapists tended to underestimate the bond relative to their clients.

---

[2] Running a model in which the truth was also treated as random at Level 3 did not improve the model fit, as indicated by a deviance test; $\chi^2(2) = 3.6$, *ns*.

Table 2
The Truth-and-Bias Model for Alliance Ratings

Parameter estimates	Estimate (SE)	p value	Effect size[a]
Fixed effects			
Intercept (discrepancy) (γ_{000})	−0.53 (0.07)***	<0.001	0.52
Slope (congruence) (γ_{100})	0.39 (0.03)***	<0.001	0.47
Random effects			
Level 1 (sessions)			
Residual	0.47 (0.01)***	<0.001	
Level 2 (clients)			
Intercept	0.33 (0.04)***	<0.001	
Slope	0.07 (0.02)***	<0.001	
Intercept/slope covariance	−0.05 (0.02)*	0.020	
Level 3 (therapists)			
Intercept	0.13 (0.04)**	0.002	
Model summary			
−2 LogL (deviance)	7038.6		
No. estimated parameters	7		

[a] Effect sizes were estimated with semipartial R^2 for linear mixed models (Edwards et al., 2008).
* $p < .05$. ** $p < .01$. *** $p < .001$.

Despite this variability, the client-specific directional bias for most therapists (78.9%) was negative, indicating that, with few exceptions, underestimation was indeed the standard (see Figure 2).

The fixed effect for the truth force was positive and significant, indicating that, as predicted (Hypothesis 1b), therapists' bond ratings were temporally congruent with their clients' bond ratings session by session. The random effect of the truth force was also significant at Level 2, indicating significant between-client variability in congruence. Despite this variability, the client-specific truth force for almost all therapists (99.5%) was positive, indicating that congruence was indeed the standard (see Figure 2).

To rule out the possibility that the positive truth force reflected a positive colinear trend of time (i.e., that both clients' and therapists' ratings increased with time, as was indeed the case, which may account for the positive truth slope; see Bolger & Laurenceau, 2013), we repeated the model with time (coded as 0 at the first session) and time*truth force terms included in the model. In this model, the truth force remained significant (estimate = 0.28, SE = 0.05, p < 0.001, effect size = 0.16) and not moderated by time (estimate = 0.00, SE = 0.01, p = 0.596). The main effect of time was positive and significant (estimate = 0.04, SE = 0.004, p < 0.001, effect size = 0.30): as the treatment progressed, therapists tended to have higher ratings of the bond. Effect sizes were estimated with semipartial R^2 for linear mixed models (Edwards, Muller, Wolfinger, Qaqish, & Schabenberger, 2008).

Finally, as predicted (Hypothesis 1c), we found a negative and significant association, r = −.33, p = .020 between client-level directional biases and truth forces (see Figure 2). This finding implies that clients whose therapists' ratings tracked theirs more closely also tended to have a greater mean-level discrepancy between their and their therapist's ratings.

Moderated Models

Pretreatment symptom severity and PD diagnosis as moderators of congruence and discrepancy. In the next set of analyses, we examined whether pretreatment symptom severity or PD diagnosis moderated either the directional bias or the truth force. Because these moderators are client-level variables, they were treated as Level-2 predictors. We performed two three-level models (one for each moderator), in which directional bias was treated as random at both Level 2 (i.e., the client level) and Level 3 (i.e., the therapist level); the truth variable (i.e., clients' ratings) was treated as random at Level 2; and pretreatment symptom severity or PD as well as their interaction with the truth force were treated as fixed effects.[3] The model's equations were:

Level 1:

$$Judgment_{sct} = \pi_{0ct} + \pi_{1ct} * Truth_{sct} + e_{sct}$$

Level 2:

$$\pi_{0ct} = \beta_{00t} + \beta_{01} * Moderator_{ct} + r_{0ct}; \pi_{1ct} = \beta_{10t}$$
$$+ \beta_{11} * Moderator_{ct} + r_{1ct}$$

Level 3:

$$\beta_{00t} = \gamma_{000} + u_{00t}; \beta_{01t} = \gamma_{010}; \beta_{10t} = \gamma_{100}; \beta_{11t} = \gamma_{110}.$$

Contrary to our predictions (Hypotheses 2a, 2b, 2c, and 2d), for both moderators, the analyses revealed neither a significant main effect (for pretreatment BSI: estimate = −0.04, SE = 0.07, p = 0.585; for PD diagnosis: estimate = −0.02, SE = 0.12, p = .889) nor a significant interaction with the truth force (for pretreatment BSI: estimate = −0.02, SE = 0.05, p = .611; for PD diagnosis: estimate = −0.02, SE = 0.08, p = .834).

Session-level symptom severity as a moderator of congruence and discrepancy. Next, we tested session-by-session symptom severity as a moderator of congruence and discrepancy (*exploratory* Hypotheses 2e and 2f). As this was a session-level variable it was treated as a Level-1 moderator. We performed a three-level model in which directional bias was treated as random at both Level 2 (i.e., the client level) and Level 3 (i.e., the therapist level); the truth variable (i.e., clients' ratings) was treated as random at Level 2; and session-by-session symptom severity as well as its interaction with the truth force were treated as random at Level 2.[4] The model's equations were:

Level 1:

$$Judgment_{sct} = \pi_{0ct}$$
$$+ \pi_{1ct} * Truth_{sct} + \pi_{2ct} * Symp_{sct} + \pi_{3ct} * Truth_{sct} * Symp_{sct} + e_{sct}$$

Level 2:

$$\pi_{0ct} = \beta_{00t} + r_{0ct}; \pi_{1ct} = \beta_{10t} + r_{1ct};$$
$$\pi_{2ct} = \beta_{20t} + r_{2ct}; \pi_{3ct} = \beta_{30t} + r_{3ct}$$

Level 3:

$$\beta_{00t} = \gamma_{000} + u_{00t}; \beta_{10t} = \gamma_{100}; \beta_{20t} = \gamma_{200}; \beta_{30t} = \gamma_{300}.$$

---

[3] Treating the main effects of the moderators and their interaction with the truth force as random at Level 3 did not improve the model's fit, as indicated by a deviance test; $\chi^2(3) = 5.2$, *ns* for pretreatment symptoms severity; $\chi^2(3) = 0.20$, *ns* for PD diagnosis. Therefore, we used the model described above. See the online supplementary Appendix for fuller tables of the moderated models.

[4] Treating the moderator and its interaction with the truth force as random at Level 3 did not improve the model's fit, as indicated by a deviance test; $\chi^2(2) = 0.9$, *ns* Therefore, we used the model described above.
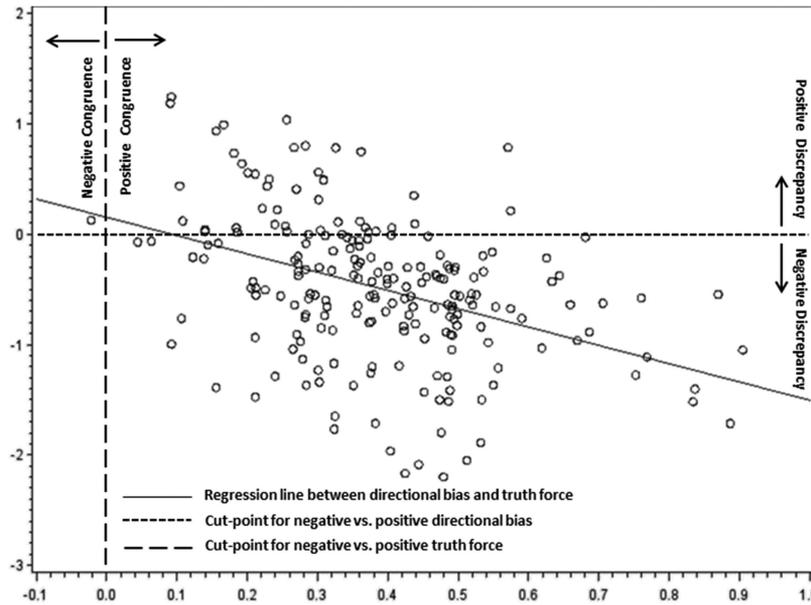
Figure 2. Therapist-client dyads' discrepancy (directional-biases), congruence (truth forces), and the associations between them.

The moderator's (fixed) effect on the therapists' judgments was negative and significant (estimate = -0.19, $SE = 0.05$, $p < .001$, effect size = 0.10). The interaction between the moderator and the truth force was also significant (estimate = -0.20, $SE = 0.08$, $p = .020$, effect size = 0.23). To examine the simple effects of the directional bias and the truth force at various levels of symptoms (i.e., low [-1 SD], average, and high [1 SD]), we used Preacher, Curran, and Bauer's (2006) computational tool for probing inter-

action effects in MLM analyses. Figure 3 displays the simple effects at the different symptom levels and shows that the directional bias became more negative and the truth force decreased as clients' symptoms increased. These findings indicate that in sessions in which clients reported higher symptomatology, therapists tended to rate the bond even lower than, *and* be less congruent with, their clients. At all three symptom levels, the simple effects of both the directional bias and the truth force remained significant.

Discussion

This study was motivated by a puzzling paradox in the literature regarding congruence between clients' and therapists' judgments of alliance—namely, the evidence for both congruence (Marmarosh & Kivlighan, 2012) and discrepancy (Fitzpatrick et al., 2005) in clients' and therapists' ratings of the alliance. We utilized West and Kenny's T&B model, an analytic framework specifically developed for these purposes, which allowed us to determine whether clients and therapists ratings were congruent, discrepant, or both.

As a prerequisite for our analyses, we first ensured that both clients' and therapists' bond ratings fluctuated significantly from session to session, as has been reported previously by multiple studies (e.g., Bachelor & Salame, 2000). That was indeed the case. Consistent with our *primary* Hypothesis (1a), we found that therapists tend to rate the bond as lower than their clients (negative mean-level directional bias) across the sessions. This is consistent with work documenting therapists' tendency toward underestimation (Fitzpatrick et al., 2005; Tryon et al., 2007). However, whereas previous research has been based on single-session or pooled alliance ratings, the current study suggests that underestimation also occurs on session-by-session.

Our analyses also supported our second *primary* Hypothesis (1b), that therapists' bond ratings would be substantially congruent with their clients' bond ratings across sessions. The obtained congruence is
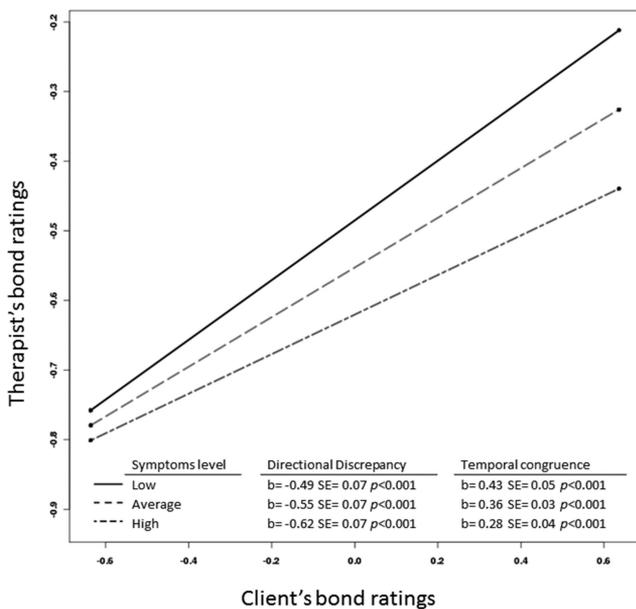

Figure 3. Directional discrepancy and temporal congruence estimates at various levels of session-level symptom severity.

consistent with previous studies that have found a moderate correlation between clients' and therapists' ratings of the alliance at the between-dyad level (cf., Tryon et al., 2007). However, it again extends prior work by showing that congruence was present at the within-dyad level, being the first to demonstrate this finding with session-by-session data.

Heeding the suggestions of Bolger and Laurenceau (2013), we wanted to ensure that the congruence found was not an artifact of the time that had elapsed in the therapy, which could potentially account for both clients' and therapists' increased bond ratings and, therefore, reflect a simple colinear trend. This examination indicated that time indeed played a role in predicting both parties' bond ratings: as the treatment progressed, therapists and clients tended to have higher bond ratings. Additionally, time moderated the directional bias (found in Hypothesis 1a): Although therapists' ratings remained lower than their clients' ratings as the treatment progressed, the discrepancy between them decreased over time. These findings echo those of Kivlighan and Shaughnessy (1995), who found evidence for lower discrepancy between clients' and therapists' mean alliance ratings in later sessions compared with earlier sessions (though see Fitzpatrick et al., 2005). Another finding of Kivlighan and Shaughnessy (1995) was that over time, clients and therapists come to share more common perceptions as reflected in a considerable rise in the correlations between their ratings, as sessions progress. In our data (Hypothesis 1b), congruence was found not to be moderated by time. For our purpose, this indicates that congruence is not a simple artifact of time. Still, it seems to conflict with Kivlighan and Shaughnessy's findings. Notably, however, their definition of congruence involved repeated sample-level correlations, whereas ours involve dyad-level correlations. These divergent results warrant further study.

West and Kenny's (2011) T&B model also allowed us to examine the degree to which temporal congruence and directional discrepancy were related, which was a critical lacuna in the existing alliance literature. Our results supported our third *primary Hypothesis* (*1c*): therapeutic dyads in which the therapist/client discrepancy was larger, were also characterized by more congruence.

The three primary findings regarding negative directional bias, temporal congruence, and the association between the two can be understood as reflecting a better safe than sorry pattern similar to the one found to characterize judgments in close dyadic relationships outside of psychotherapy (Fletcher & Kerr, 2010; Overall et al., 2012). Exploring close dyadic relationships, Overall and her colleagues focused on perceivers' inferences about their partners' thoughts and feelings that reflected the connection between the self and the partner. In this context, they found that a combined pattern of negative directional bias and temporal congruence was most adaptive, and argued that this pattern helps avoid the risk of overestimating a partner's goodwill and failing to detect problems in the relationship.

As Marmarosh and Kivlighan (2012) recently posited, a similar pattern may be present in therapeutic relationships, in which overestimation might cause missed potential ruptures (cf., Safran & Muran, 1996). Specifically, these authors found that clients experienced sessions as smoother if their therapists' estimation of the alliance was lower than their own. Our results build upon these ideas, suggesting that therapists' vigilance may manifest itself in both a negative bias and a temporal congruence, and that the two are tied together. The same risks that produce cautious underestimation of the bond, also seems to attune therapists to fluctuations in it.

Earlier treatments of the underestimation issue have offered alternative explanations for this phenomenon. One theory focused on therapists' ability to compare a particular therapy relationship to others, an ability not shared by their clients (Fenton et al., 2001). An additional theory (e.g., Bachelor & Salame, 2000; Fitzpatrick et al., 2005) noted that most therapists in psychotherapy studies are trainees who tend to be unsure of their competence and, therefore, underestimate the quality of their work (and the bond). Yet, another possibility is that clients provide high ratings, either because of the social desirability of showing appreciation for their therapist, or because of the motivation of reducing the cognitive dissonance generated by the voluntary attendance of therapy; unlike the clients, therapists might lack these motivations. More important, the better safe than sorry explanation is not inconsistent with any of these alternative explanations, but offers a parsimonious account for all three of our findings (i.e., underestimation, temporal congruence, and a positive association between them). It is also not inconsistent with a counterpart explanation, which would be that the (satisfied) clients of more congruent therapists rate the therapeutic bond more positively, thus increasing the directional bias.

Our study's second broad goal was to examine whether client distress moderates discrepancy or temporal congruence between therapists' and clients' bond ratings. Although clients' distress have previously been found to moderate the sample level congruence between clients' and therapists' alliance ratings (Tryon et al., 2007), they have not previously been studied as moderators of congruence or discrepancy at the person or the session level.

At the person level, we expected PD diagnoses and pretreatment symptom severity to be associated with more moderate underestimation bias (Hypotheses 2a and 2b). We also expected PD symptoms to be associated with lower temporal congruence (Hypothesis 2c). Finally, we examined (without prior expectation) the association between symptom severity and temporal congruence (Hypothesis 2d). None of these associations, predicted or otherwise, proved significant. Specifically, preexisting symptomatology or PD diagnoses did not moderate discrepancy or temporal congruence. Because this investigation is the first to examine these moderators vis-à-vis bond congruence and the first to use session-by-session data, these associations should be further examined.

We also explored the associations between temporal congruence and directional discrepancy and symptoms at the session level. Both the directional discrepancy and the congruence decreased in sessions in which clients' symptoms increased. In these sessions, therapists tended to be even more discrepant from their clients, and to be less congruent with their clients' experience of the bond. A possible explanation for this finding is that in sessions in which the clients' symptoms worsened, therapists tended to incorrectly attribute the clients' behavior to (rifts in) the therapeutic bond and, therefore, assumed that the bond was damaged, leading to the increased directional bias. This misattribution may have also led therapists to be less attuned to in-session processes that affect clients' experience of the bond, resulting in decreased congruence. Indeed, as Markin, Kivlighan, Gelso, Hummel and Spiegel, (2014) recently noted, therapists and clients may base their judgments (in this case, of the bond) on somewhat different factors. Alternatively, therapists may respond to clients' greater symptomatology with their own upset; they may then become distracted, pessimistic, or anxious, and as a consequence, be both less attentive to in-session events that impact the bond *and* provide lower bond ratings overall. Significantly, even in high-

distress sessions (i.e., at 1 *SD* above the average symptom score), temporal congruence remained significantly positive (if weaker), suggesting that therapists retained some level of attunement.

Limitations, Future Directions, and Summary

One limitation of this study was that alliance was indexed using the global alliance subscale of the BPSR (Flückiger et al., 2010). Though this scale does cohere quite closely to Bordin's (1994) *bond* concept, future studies may benefit from examining congruence between therapists' and clients' ratings of the other facets of the therapeutic alliance (namely, *tasks* and *goals*). Additionally, the BPSR asks both clients and therapists to provide (independent) judgments of the bond; it would be interesting to apply the T&B model to data in which the judgment variable (i.e., the therapists' ratings) explicitly involves the therapists' inferences regarding the clients' bond perceptions (i.e., in the case in which the clients' ratings are indeed the "truth" criterion). Doing so may call for using strictly parallel scales, such as the HAQ's (Luborsky et al., 1996). At the same time, the BPSR used here has the advantage of being brief and, therefore, more appropriate for repeated session-by-session administration.

Therapist characteristics (e.g., interpersonal skills, professional experience) that may affect accuracy or bias were not assessed in this study. Our results indicated that only 16.3% of the variance in therapists' bond ratings, and much less in the clients' ratings, were attributable to the therapist level. Of interest to the authors, other authors have found that therapists' characteristics account for more substantial variance in their judgments (e.g., 38% of the variance in transference ratings; Markin & Kivlighan, 2007) and that between-therapist variability plays a much larger part than between-client variability in the association of alliance and outcome (Del Re et al., 2012). It is hard to know why our therapists were less variable, and future studies should explore this issue of between-therapist variability in more depth. It is possible that the adherence of all study therapists to one orientation, namely CBT, and the reliance on trainee therapists was partly responsible for this homogeneity; this may limit the generalizability of the present findings to therapies following other orientations or those implemented by more experienced clinicians.

The prevalence of one of our moderators, PD diagnosis, was relatively low (19%) compared with rates in other outpatient samples (45.5%; Zimmerman, Rothschild, & Chelminski, 2005). The data reported here were collected in a training clinic that screens out many clients with more complicated diagnostic profiles; thus, the results need to be replicated in other samples with more typical rate of PD diagnosis.

Finally, this study utilized session-by-session data from the initial months of therapy. It would be important to reexamine the finding that time in treatment moderated directional bias but did not moderate tracking accuracy with data spanning additional periods within the therapy or even the entire course of therapy. For example, the discrepancy between therapists' and clients' ratings of the bond may continue to decrease over the course of the entire treatment.

These limitations notwithstanding, the present study extends the examination of congruence between clients' and therapists' bond perceptions in several ways. To our knowledge, this research is the first study to use session-by-session data to assess congruence and discrepancy in clients' and therapists' bond ratings. Using the novel statistical procedure developed by West and Kenny (2011), we simultaneously assessed temporal congruence and mean-level discrepancy

and demonstrated that the two co-occur in what we call a better safe than sorry pattern.

The study is also innovative in examining possible moderators of bias and accuracy. Of interest to the authors, we did not find preexisting client characteristics (i.e., symptom severity or PD diagnosis) to moderate the extent to which congruence or discrepancy occurred across the sessions. However, we did find that session-level symptom severity was associated with greater therapist underestimation as well as reduced temporal congruence. We believe that these findings reflect therapists' misattribution of clients' symptoms to the therapeutic relationship.

Our results have several possible clinical implications. First, they advance the idea that therapists who adopt a vigilant (and somewhat pessimistic) approach may be more attuned to their clients' changing experience. Second, the findings highlight the risk of misattributing symptomatic change to factors within the therapeutic relationship. Future work should further explore both of these possibilities. This research can be aided by the availability of the T&B model, a novel methodological approach in the field of psychotherapy research.

References

Bachelor, A., & Salamé, R. (2000). Participants' perceptions of dimensions of the therapeutic alliance over the course of therapy. *The Journal of Psychotherapy Practice and Research, 9,* 39–53.

Bolger, N., & Laurenceau, J. P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York, NY: Guilford Press.

Bordin, E. S. (1994). Theory and research on the therapeutic working alliance: New directions. In A. O. Horvath & L. S. Greenberg (Eds.), *The working alliance: Theory, research, and practice* (pp. 13–37). New York, NY: Wiley.

Bronisch, T., Hiller, W., Mombour, W., & Zaudig, M. (1996). *International diagnostic checklists for personality disorders according to ICD-10 and DSM–IV-IDCL-P*. Seattle, WA: Hogrefe and Huber Publishers.

Connolly Gibbons, M. B., Crits-Christoph, P., de la Cruz, C., Barber, J. P., Siqueland, L., & Gladis, M. (2003). Pretreatment expectations, interpersonal functioning, and symptoms in the prediction of the therapeutic alliance across supportive-expressive psychotherapy and cognitive therapy. *Psychotherapy Research, 13,* 59–76. http://dx.doi.org/10.1093/ptr/kpg007

Del Re, A. C., Flückiger, C., Horvath, A. O., Symonds, D., & Wampold, B. E. (2012). Therapist effects in the therapeutic alliance-outcome relationship: A restricted-maximum likelihood meta-analysis. *Clinical Psychology Review, 32,* 642–649. http://dx.doi.org/10.1016/j.cpr.2012.07.002

Derogatis, L. R. (1975). *Brief symptom inventory*. Baltimore, MD: Clinical Psychometric Res.

Derogatis, L. R. (1992). *The symptom checklist-90-revised*. Minneapolis, MN: NCS.

Edwards, L. J., Muller, K. E., Wolfinger, R. D., Qaqish, B. F., & Schabenberger, O. (2008). An R^2 statistic for fixed effects in the linear mixed model. *Statistics in Medicine, 27,* 6137–6157. http://dx.doi.org/10.1002/sim.3429

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

Fletcher, G. J. O., & Kerr, P. S. G. (2010). Through the eyes of love: Reality and illusion in intimate relationships. *Psychological Bulletin, 136,* 627–658. http://dx.doi.org/10.1037/a0019792

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,* 1–13. http://dx.doi.org/10.1080/10503307.2012.693773

Flückiger, C., Regli, D., Zwahlen, D., Hostettler, S., & Caspar, F. (2010). Der Berner Patienten- und Therapeutenstundenbogen 2000. *Zeitschrift für Klinische Psychologie und Psychotherapie, 39,* 71–79. http://dx.doi.org/10.1026/1616-3443/a000015

Franke, G. (2000). *BSI. Brief symptom inventory: Deutsche version. Manual.* Göttingen: Beltz.

Grawe, K. (1997). Research-informed psychotherapy. *Psychotherapy Research, 7,* 1–19. http://dx.doi.org/10.1080/10503309712331331843

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

Hartmann, A., Joos, A., Orlinsky, D. E., & Zeeck, A. (2014). Accuracy of therapist perceptions of patients' alliance: Exploring the divergence. *Psychotherapy Research, 24,* 1–12. http://dx.doi.org/10.1080/10503307.2014.927601

Haselton, M. G., & Buss, D. M. (2000). Error management theory: A new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology, 78,* 81–91. http://dx.doi.org/10.1037/0022-3514.78.1.81

Hersoug, A. G., Høglend, P., Havik, O. E., & Monsen, J. T. (2010). Development of working alliance over the course of psychotherapy. *Psychology and Psychotherapy: Theory, Research and Practice, 83,* 145–159. http://dx.doi.org/10.1348/147608309X471497

Hilsenroth, M. J., Peters, E. J., & Ackerman, S. J. (2004). The development of therapeutic alliance during psychological assessment: Patient and therapist perspectives across treatment. *Journal of Personality Assessment, 83,* 332–344. http://dx.doi.org/10.1207/s15327752jpa8303_14

Horvath, A. O., Del Re, A. C., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy, 48,* 9–16. http://dx.doi.org/10.1037/a0022186

Hox, J. (2010). *Multilevel analysis: Techniques and applications.* England: Routledge.

Kenny, D. A. (2004). PERSON: A general model of interpersonal perception. *Personality and Social Psychology Review, 8,* 265–280. http://dx.doi.org/10.1207/s15327957pspr0803_3

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics, 53,* 983–997.

Kivlighan, D. M., & Shaughnessy, P. (1995). Analysis of the development of the working alliance using hierarchical linear modeling. *Journal of Counseling Psychology, 42,* 338–349. http://dx.doi.org/10.1037/0022-0167.42.3.338

Levy, K. N., Beeney, J. E., Wasserman, R. H., & Clarkin, J. F. (2010). Conflict begets conflict: Executive control, mental state vacillations, and the therapeutic alliance in treatment of borderline personality disorder. *Psychotherapy Research, 20,* 413–422. http://dx.doi.org/10.1080/10503301003636696

Luborsky, L., Barber, J. P., Siqueland, L., Johnson, S., Najavits, L. M., Frank, A., & Daley, D. (1996). The revised helping alliance questionnaire (HAq-II): Psychometric properties. *The Journal of Psychotherapy Practice and Research, 5,* 260–271.

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. *Diagnostica, 52,* 11–25. http://dx.doi.org/10.1026/0012-1924.52.1.11

Markin, R. D., & Kivlighan, D. M. (2007). Bias in psychotherapist ratings of client transference and insight. *Psychotherapy, 44,* 300–315. http://dx.doi.org/10.1037/0033-3204.44.3.300

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

Muran, J. C., Safran, J. D., Gorman, B. S., Samstag, L. W., Eubanks-Carter, C., & Winston, A. (2009). The relationship of early alliance ruptures and their resolution to process and outcome in three time-limited psychotherapies for personality disorders. *Psychotherapy, 46,* 233–248. http://dx.doi.org/10.1037/a0016085

Orlinsky, D. E., & Howard, K. I. (1986). Process and outcome in psychotherapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed., pp. 311–381). New York, NY: Wiley.

Overall, N. C., Fletcher, G. J. O., & Kenny, D. A. (2012). When bias and insecurity promote accuracy: Mean-level bias and tracking accuracy in couples' conflict discussions. *Personality and Social Psychology Bulletin, 38,* 642–655. http://dx.doi.org/10.1177/0146167211432764

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics, 31,* 437–448. http://dx.doi.org/10.3102/10769986031004437

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

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

Tschacher, W., Ramseyer, F., & Grawe, K. (2007). Der ordnungseffekt im psy- chotherapieprozess: Replikation einer systemtheoretischen vorhersage und zusam- menhangmitdem therapieerfolg [The order effect in the psychotherapy process]. *Zeitschrift für Klinische Psychologie und Psychotherapie, 36,* 18–25. http://dx.doi.org/10.1026/1616-3443.36.1.18

Tufekcioglu, S., Muran, J. C., Safran, J. D., & Winston, A. (2013). Personality disorder and early therapeutic alliance in two time-limited therapies. *Psychotherapy Research, 23,* 646–657. http://dx.doi.org/10.1080/10503307.2013.843803

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

Zimmerman, M., Rothschild, L., & Chelminski, I. (2005). The prevalence of *DSM–IV* personality disorders in psychiatric outpatients. *The American Journal of Psychiatry, 162,* 1911–1918. http://dx.doi.org/10.1176/appi.ajp.162.10.1911