American Psychological Association

Journal of Consulting and Clinical Psychology

0022-006X/19/$12.00

2019, Vol. 87, No. 1, 33–45
http://dx.doi.org/10.1037/ccp0000354

Therapists' Empathic Accuracy Toward Their Clients' Emotions

Dana Atzil-Slonim
Bar-Ilan University

Eran Bar-Kalifa
Ben-Gurion University of the Negev

Hadar Fisher, Gal Lazarus, and Ilanit Hasson-Ohayon
Bar-Ilan University

Wolfgang Lutz and Julian Rubel
University of Trier

Eshkol Rafaeli
Bar-Ilan University

Objective: Therapists' empathic accuracy (EA) toward their clients' fluctuating emotions is a crucial clinical skill that underlies many therapeutic interventions. In contrast to the subjective components of empathy, limited empirical work has addressed EA or its effect on the outcomes of psychotherapy. Here, we differentiate between the components of EA (tracking accuracy, directional discrepancy) as well as the valence of the target emotions (positive vs. negative). We also investigated the relative contribution of cognitive and emotional processes to therapists' EA and examined the associations between EA and treatment outcomes. ***Method:*** The sample comprised 93 clients treated by 62 therapists in a university setting. Prior to each session, clients self-reported their symptoms. Following each session, clients rated their positive (PE) and negative (NE) emotions during the session and therapists rated their own emotions, as well as their assessment of their clients' emotions. ***Results:*** Therapists accurately tracked their clients' PE and NE and were more accurate for NE. Therapists tended to overestimate their clients' NE and underestimate their clients' PE. Therapists' emotions were associated with their clients' emotions (real similarity). Therapists' emotions were also associated with their assessments of their clients' emotions (assumed similarity). Therapists' own emotions partially mediated the association between clients' emotions and therapists' assessments. Therapists' inaccuracy in assessing their clients' PE was associated with higher reported symptoms in the next session. ***Conclusion:*** These findings help provide a better understanding of the specific characteristics associated with more EA and underscore the importance of EA in facilitating clients' emotional well-being.

What is the public health significance of this article?
The current findings highlight the importance of therapists' empathic accuracy regarding their clients' emotions. The findings advance the idea that both cognitive and emotional empathy contribute to therapists' ability to correctly assess their clients' emotions. The results point to the risk on the part of therapists to neglect clients' positive emotions and stress the importance of attending to these emotions.

Keywords: empathic accuracy, empathy, emotions, truth and bias model, polynomial regression with response surface analysis

The ability to accurately perceive and assess the mental states of other individuals plays a crucial role in most interpersonal relationships (for a review, see Hall, Mast, & West, 2016). This ability, often referred to as *empathic accuracy* (EA; Ickes & Hodges, 2013), has been investigated through multiple paradigms and appears to be associated with positive outcomes in various types of relationships (e.g., Rafaeli, Gadassi, Howland, Boussi, & Lazarus, 2017; Sened, Yovel, Bar-Kalifa, Gadassi, & Rafaeli, 2017).

EA may be considered one variant of empathy, a broad term that encompasses a range of facets of interpersonal attunement and sensitivity (cf. Decety & Ickes, 2011). The concept of empathy was introduced to the field of psychotherapy by Rogers (1957), who defined it as the ability to accurately perceive the internal frames of reference of other individuals and posited that it constituted a necessary condition for effective psychotherapy. Psychotherapy researchers and practitioners have shown substantial interest in empathy and view it as both a skill and an attitude (e.g., Barrett-Lennard, 1981; Greenberg & Elliott, 1997; Kohut, 1977;

This article was published Online First November 26, 2018.

Dana Atzil-Slonim, Department of Psychology, Bar-Ilan University; Eran Bar-Kalifa, Department of Psychology, Ben-Gurion University of the Negev; Hadar Fisher, Gal Lazarus, and Ilanit Hasson-Ohayon, Department of Psychology, Bar-Ilan University; Wolfgang Lutz and Julian Rubel, Department of Psychology, University of Trier; Eshkol Rafaeli, 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, Department of Psychology, Bar-Ilan University, Ramat-Gan, 5290002 Israel. E-mail: dana.slonim@gmail.com

Thwaites & Bennett-Levy, 2007); it has also been used to predict treatment outcomes across many therapeutic modalities (cf. Elliott, Bohart, Watson, & Greenberg, 2011; Watson, 2016). Some aspects of empathy and particularly its subjective experience have received considerable attention within psychotherapy research, whereas other aspects, including the objective assessment of the accuracy of empathic inferences, have been left relatively unexplored (Elliott et al., 2011).

Therapists' ability to be empathically accurate is particularly important with regard to clients' emotions because being correctly attuned to changes in clients' fluctuating emotions underlies many therapeutic interventions (Elliott et al., 2011; Watson, 2016). Numerous studies have shown that clients' experiences of emotions during psychotherapy represent a strong predictor of positive psychotherapeutic outcomes across treatment modalities (for reviews, refer to Greenberg, 2010; Hayes, Ready, & Yasinski, 2014; Hofmann & Hayes, 2015). Several studies have suggested that therapists' emotions are also an important factor in treatment outcomes (e.g., Hayes, Gelso, & Hummel, 2011; Holmqvist, Hansjons-Gustafsson, & Gustafsson, 2002; Westra, Aviram, Connors, Kertes, & Ahmed, 2012). However, despite the growing consensus among psychotherapy theorists and researchers that emotions should be studied as dynamic systems that interact over time not only within the client or therapist (i.e., intrapersonally) but also between the two individuals (i.e., interpersonally; cf. Fosha, 2001; McCullough et al., 2003), surprisingly little research has addressed how therapists' own emotions interact with the emotions of their clients during psychotherapy (for exceptions, refer to Chui, Hill, Kline, Kuo, & Mohr, 2016; Duan & Kivlighan, 2002). The importance of investigating emotions both intrapersonally and interpersonally is particularly germane to the study of empathy, given its inherently relational nature (Main, Walle, Kho, & Halpern, 2017).

In affective neuroscience studies there is a growing consensus that empathy consists of two major processes: a cognitive one and an emotional one (cf. Cuff, Brown, Taylor, & Howat, 2016; Shamay-Tsoory, 2011). The cognitive process (also referred to as the mental state attribution system or simply as mentalizing; Zaki & Ochsner, 2012) involves (successful or unsuccessful) understanding of another individual's mental states and includes a more deliberate processing of cues. In psychotherapy, this would imply that the therapist takes the perspective of the client, which helps her accurately understand what the client feels. The emotional process (also referred to as the experience sharing system, or simply as emotional resonance; Zaki & Ochsner, 2012) is driven by a more automatic sharing of mental states. In psychotherapy, this would imply that the therapist experiences (at least some of) the same emotions as the client (Duan & Kivlighan, 2002), and uses this shared experience as a source of information when inferring the client's emotions. Figure 1 depicts how both processes can culminate in EA.

To see this intuitively, consider a hypothetical therapeutic dyad: Louis the client and Hannah the therapist. Louis may feel sadness during the session as he touches on his painful relationship with his father. Hannah may cognitively process the verbal and nonverbal information that she perceives from Louis and be empathically accurate when she assesses that he feels sad. This cognitive process may be considered a direct accuracy path (see path c′ in Figure 1), as it involves a single deliberate step of cue-processing

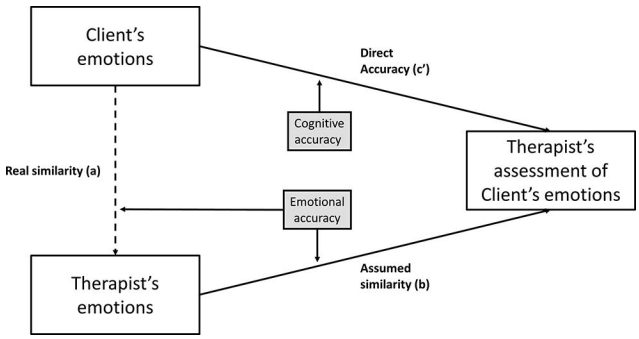

Figure 1. Illustration of the two paths to empathic accuracy (EA): a (real similarity) × b (assumed similarity) is the emotional path to accuracy; c′ (direct accuracy) is the cognitive path to accuracy. Note that the association between clients' emotions and therapists' emotions is marked as a dashed line to highlight that it does not represent a predictive association (with clients' emotions preceding therapists' emotions) but rather a correlational path (i.e., clients' emotions are associated with therapists' emotions).

to assess another individual's state of mind (e.g., Kenny & Acitelli, 2001).

Alongside this cognitive empathy, Hannah may also experience sadness as she listens to Louis and may use her own feeling, explicitly or implicitly, to (accurately) conclude that Louis feels the same. This emotional process may be considered an indirect accuracy path, as it involves two steps to arrive at accurate assessments—*real similarity* (the link between Hannah's and Louis' emotions, see path a in Figure 1) and assumed similarity (the link between Hannah's emotions and her judgment of Louis' emotions, see path b in Figure 1). *Assumed similarity*, the assumption that the other feels like the perceiver, has been discussed as a form of projective bias on the part of the perceiver (in this case, Hannah; e.g., Kenny & Acitelli, 2001; Overall, Fletcher, Simpson, & Fillo, 2015).

In considering this model, two points are noteworthy. First, the terms *cognitive* and *emotional* are used for consistency with extant social neuroscience work (e.g., Shamay-Tsoory, 2011) do not assume that cognition and emotion are distinct systems for the target. Instead, they focus on the perceiver's experience, and distinguish between the cognitive process of mentalization and the emotional experience of resonance (Zaki & Ochsner, 2012). Of course, emotion and cognition are deeply intertwined (Ledoux & Hofmann, 2018; Pessoa, 2008; Preston & Hofelich, 2012).

Second, the terms *direct* and *indirect*, also used for consistency with previous research (e.g., Kenny & Acitelli, 2001), do not refer to the immediacy of the accuracy experience or lack thereof. In fact, the indirect path may well be more automatic, and less conscious or deliberate, than the direct path (Cuff et al., 2016; Hodges & Wegner, 1997). Its indirectness simply refers to the way in which it involves two steps rather than one.

When Hannah accurately assesses Louis's emotions, does her accuracy require cognitive processing? Does it benefit from emotional similarity, or instead, does this similarity impede accuracy? Both the cognitive and emotional processes to accuracy have been extensively investigated in relationship science (e.g., Sened et al., 2017; Wilhelm & Perrez, 2004). In contrast, in psychotherapy research, despite repeated calls (e.g., Bohart & Greenberg, 1997;

Elliott et al., 2011; Watson, 2016), only one study to date has done so (Duan & Kivlighan, 2002). This study differentiated between the *cognitive process* (defined as the similarity between clients' perception of their own emotions and therapists' perception of these emotions) and the *emotional process* (defined as the similarity between the clients' emotions and the therapists' emotions) to therapist EA as assessed during one session in the middle of therapy. Both processes were found to contribute to session outcomes.

In the current study, we monitored clients' emotions, therapists' emotions, and therapists' assessments of their clients' emotions on a session-by-session basis over the course of treatment. This enabled us to assess the relative contribution of the cognitive and emotional processes that are associated with therapists' EA regarding their clients' emotions. It also enabled us to test the association between this accuracy and treatment outcomes. Thus, we extend the work of Duan and Kivlighan (2002) in several ways. In particular, these authors relied on similarity/distance indices and only assessed emotional and cognitive empathy once during treatment. In contrast, the current data were obtained by tracking session-by-session fluctuations in clients' and therapists' emotions, as well as therapists' assessments of the clients' emotions.

Methods for Assessing EA

The paucity of studies examining EA within psychotherapy may have to do, at least in part, with the overreliance of the field on subjective judgments of empathy (from the client, therapist, or observer perspective; Watson, 2016) and the underuse of objective measures of accuracy (for a review, see Elliott et al., 2011).

Outside the clinical realm, one of the most frequently used methods for EA assessment has been the dyadic interaction paradigm developed by Ickes (2003). In this paradigm, dyads are recorded during a videotaped interaction, and each partner subsequently reviews the recording, recalls his or her own feelings and infers the partner's feelings. Objective observers then rate the correspondence between the targets' recollections and the perceivers' inferences. Several studies have used this method or variants in clinically relevant settings. In two such studies, students viewed recordings of real (Barone et al., 2005) or simulated (Marangoni, Garcia, Ickes, & Teng, 1995) therapy sessions and inferred the clients' feelings. In contrast, Kwon and Jo (2012) used this method in a real-life psychotherapy setting and instructed clients and therapists to review the recordings of their first three sessions and provide their actual (or inferred) feelings. The correspondence between these actual and inferred feelings served as the index of EA and was associated with treatment outcome. In short, despite its merits (e.g., high internal validity), the dyadic interaction paradigm is both difficult to implement and limited in terms of its external validity in the context of psychotherapy.

An emerging alternative method for EA assessment employs ecologically valid repeated assessments (e.g., diary data) within naturalistic settings (Howland & Rafaeli, 2010). Both the perceiver (e.g., the therapist) and the target (e.g., the client) are asked to rate specific variables (e.g., emotions) repeatedly; the congruence between these reports is the measure of EA. If kept brief, these assessments may be implemented repeatedly over a greater time span; for example, on a session-by-session basis over the span of the entire treatment. This approach enables researchers to examine how empathic accuracy changes throughout treatment and whether it leads to positive outcomes in psychotherapy. If the therapist's own session-by-session emotions are also assessed, this method also provides a way to examine assumed similarity.

The repeated assessment approach to operationalizing EA has two additional benefits: it enables examination of the type of accuracy, and it makes it possible to pinpoint what the therapist is accurate about and, in particular, the valence of the reported and inferred emotions. With respect to the type of accuracy, various authors have highlighted the importance of examining tracking accuracy, directional discrepancy, or both factors (e.g., Fletcher & Kerr, 2010). Tracking accuracy is defined as the correlation between a set of judgments over time (e.g., the correlation between clients' and therapists' ratings of the clients' emotions). Directional discrepancy captures the differences in the mean levels of a specific judgment across a sample compared with a benchmark (e.g., the degree to which therapists over- or underestimate their clients' emotions compared with the clients' own ratings). The distinction between these components has received considerable attention outside the clinical domain (e.g., Overall, Fletcher, & Kenny, 2012). However, until recently, its importance has been largely unrecognized in the psychotherapy literature. Recent research has begun to implement conceptually and psychometrically sound analytic approaches, such as the truth and bias model (T&B; West & Kenny, 2011), which make it possible to differentiate these two components of EA (or related constructs, such as agreement or congruence) and examine their simultaneous effects within psychotherapy (e.g., Atzil-Slonim et al., 2015; Rubel, Bar-Kalifa, Atzil-Slonim, Schmidt, & Lutz, 2018). To date, these studies have shown that therapists' accuracy in detecting their clients' experience of alliance or functioning is related to treatment outcome; however, no studies have addressed tracking accuracy or directional discrepancy with regard to emotions.

With respect to valence, although many studies of EA have not distinguished between positive emotions (PE) and negative emotions (NE; e.g., Keown & Woodward, 2002), recent work has emphasized the importance of differentiating them (Cohen, Schulz, Weiss, & Waldinger, 2012; Howland & Rafaeli, 2010). Findings outside the clinical domain have demonstrated that EA for negatively valenced states tends to be stronger than EA for positively valenced states (Howland & Rafaeli, 2010).

Within psychotherapy research, a recent study by Chui and colleagues (2016) also supported the need to distinguish between PE and NE. In another recent study (Atzil-Slonim et al., 2018), the authors used the T&B model to assess congruence between clients' and therapists' ratings of their PE and NE as they cofluctuated from session to session. Clients and therapists were temporally similar in both PE and NE. Moreover, therapists experienced less intense PE on average; however, they did not experience more or less intense NE than their clients. Finally, therapist/client congruence in both PE and NE predicted better next-session symptomatology.

Importantly, Atzil-Slonim et al. (2018), focused on congruence between clients' and therapists' ratings of their emotions (i.e., real similarity). They called for future studies assessing the extent to which therapists project their own emotions onto their clients (assumed similarity) and the extent to which therapists accurately perceive their clients' own subjective emotional experiences (empathic accuracy).

Building on this earlier work, the present study examined the role of emotional and cognitive empathic accuracy in terms of both tracking accuracy and directional discrepancy, and with both positively and negatively valenced emotions. We also examined whether EA was linked to treatment outcomes. Specifically, the following hypotheses were tested:

Hypothesis 1

On the basis of the theoretical importance of EA in psychotherapy (cf. Watson, 2016), previous research on EA within the context of psychotherapy (Kwon & Jo, 2012), and findings on EA in other types of close relationships (e.g., Sened et al., 2017), we predicted that therapists would accurately track fluctuations in their clients' PE and NE over time. We expected therapists to be more accurate regarding their clients' NE than their clients' PE, as reported in previous studies outside the clinical domain (Howland & Rafaeli, 2010).

Hypothesis 2

On the basis of previous studies that have examined congruence between clients' and therapists' emotions (Atzil-Slonim et al., 2018), as well as other process variables (e.g., Rubel et al., 2018) indicating that therapists tend to adopt a somewhat pessimistic approach toward their clients' progress, we expected that therapists would tend to underestimate their clients' PE and overestimate their clients' NE.

Hypothesis 3

On the basis of the theory regarding the importance of both cognitive and emotional processes in therapists' empathy (e.g., Watson, 2016), as well as findings on the relative contribution of cognitive and emotional accuracy (Sened et al., 2017), we expected that therapists' own self-reported emotions would partially mediate the association between clients' emotions and therapists' assessments of these emotions (for both PE and NE). For this to occur, we expected a positive association between clients' emotions and therapists' emotions (i.e., real emotional similarity; the dashed line [path a] in Figure 1) and a positive association between therapists' emotions and their assessments of their clients' emotions (i.e., assumed similarity; path b in Figure 1). We also expected that therapists' own emotions would partially mediate the association between clients' emotions and therapists' assessments (i.e., paths a × b in Figure 1). Note that our use of mediation terminology does not intend to imply a causal relationship between the variables, but rather that part of the variance in clients' emotions is shared by both the therapists' own emotions and the therapists' assessments. Thus, in purely statistical terms, and consistent with West and Kenny's (2011) original presentation of the T&B model, the therapists' own emotions can be described as a confounder rather than a mediator. Specifically, we cannot assume that therapists' emotions are necessarily caused by clients' emotions or that they lead to the therapists' assessments; instead, it may simply covary with both clients' emotions and therapists' assessments.

Hypothesis 4

On the basis of the previous studies that indicated treatment outcomes were associated with therapists' empathy (e.g., Elliott et al., 2011) and EA (Duan & Kivlighan, 2002; Kwon & Jo, 2012), we predicted that therapists' accurate perception of their clients' emotions would be associated with improvement in their clients' symptoms levels from session to session.

Method

Participants and Treatment

Clients. The analyses were based on a sample of 93 clients who were in individual psychotherapy at a large university outpatient clinic between August 2015 and August 2016. The clients included in the analysis had at least six coded sessions of individual treatment. On average, the clients received 25 treatment sessions ($SD = 9.1$). Approximately 87% ($N = 2{,}028$) of the sessions were available for analyses.

The clients were all over the age of 18 ($M_{age} = 38$ years, $SD = 13$; age range = 19–70 years), and the majority were female (56%). In the sample, 44% of the clients were single, 14% were divorced or widowed, and 42% were married or in a permanent relationship. In addition, 57% percent had at least a bachelor's degree, and 82% were fully or partially employed. The Mini-International Neuropsychiatric Interview Version 5.0 (M.I.N.I.; Sheehan et al., 1998) was used to establish an Axis I diagnosis. The interview was conducted before the actual therapy by intensively trained independent clinicians. All intake sessions were audiotaped, and a random 25% of the interviews were sampled and rated again by an independent clinician. The mean kappa value for the Axis I diagnoses was excellent ($\kappa = 0.974$).

Approximately 23% of the clients reported experiencing relationship problems, academic/occupational stress, or other problems; however, they did not meet the criteria for an Axis I diagnosis. Of the total sample, 39% had a single diagnosis, 13% had two diagnoses, and 25% had three or more diagnoses. Most clients were diagnosed with affective disorders (50%) or anxiety disorders (18%) as the primary diagnosis.[1] Additional primary diagnoses included obsessive-compulsive disorder (4%) or other disorders (5%).

Therapists. The participating clients were treated by 62 therapists (48 women and 14 men) at different stages of clinical training. The clients were assigned to the therapists in an ecologically valid manner based on real-world issues, such as therapist availability and caseload. Thirty-eight therapists treated one client each, 19 therapists treated two clients each, and five therapists treated three or more clients each. Each therapist received one hour of individual supervision and four hours of group supervision on a weekly basis. All therapy sessions were audiotaped for use in supervision. Supervisors were senior clinicians. Individual and group supervision focused heavily on the review of audiotaped case material and technical interventions designed to facilitate the appropriate use of therapists' interventions.

Individual psychotherapy consisted of once or twice weekly sessions. The dominant approach in the clinic is a short-term

[1] According to Diagnostic and statistical manual of mental disorders (*DSM–5*; American Psychiatric Association, 2013) the following diagnoses were assumed in the affective disorders cluster: major depressive disorder, dysthymia, and bipolar disorder. The following *DSM–5* diagnoses were assumed in the anxiety disorders cluster: panic disorder, agoraphobia, generalized anxiety disorder, and social anxiety disorder.

psychodynamic psychotherapy treatment model (cf. Summers & Barber, 2010). The key features of the model include the following: (a) a focus on affect and the experience and expression of emotions; (b) exploration of attempts to avoid distressing thoughts and feelings; (c) identification of recurring themes and patterns; (d) emphasis on past experiences; (e) focus on interpersonal experiences; (f) emphasis on the therapeutic relationship; and (g) exploration of wishes, dreams, or fantasies (Blagys & Hilsenroth, 2000; Shedler, 2010). Treatment was open-ended in length; however, given that psychotherapy was provided by clinical trainees at a university-based outpatient community clinic, the treatment duration was often restricted to 9 months to 1 year.

Instruments and Data Collection

Hopkins Symptom Checklist–Short Form (HSCL-11; Lutz, Tholen, Schürch, & Berking, 2006). The HSCL-11 is an 11-item inventory that is a brief version of the Hopkins Symptom Checklist-90 (SCL-90; Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974). The items are rated on a four-point Likert scale that ranges from 1 (*not at all*) to 4 (*extremely*) and refer to the previous week. Thus, the mean score of the 11 items represents the symptomatic state of the client during the previous week. It has high internal consistency ($\alpha = .92$) and concurrent validity (Lutz et al., 2006). The between- and within-person reliabilities for the scale were computed using procedures outlined by Cranford et al. (2006) for estimating reliabilities for repeated within-person measures; the reliability levels were considered high in the current study (within = 0.90; between = 0.84).

Profile of Mood States (POMS; McNair, Lorr, & Droppleman, 1992). The POMS is a widely used instrument that assesses mood variables. We used an abbreviated version of this measure, which was adapted for intensive repeated measurements (Cranford et al., 2006) and consists of 12 words that describe current emotional states. The NE items were summed to create an NE scale and the PE items were summed to create a PE scale.[2] The NE scale includes depressed mood (two items), anxious mood (two items), and anger (two items). The PE scale includes contentment (two items), vigor (two items), and calmness (two items). Examples of feelings on the POMS are "anxious," "sad," "angry," "happy," "lively," and "calm." Both clients and therapists were asked to evaluate how they felt during the session on a five-point Likert scale that ranged from 1 (*not at all*) to 5 (*extremely*). Therapists were also asked to evaluate how their clients felt during the session. The POMS has been tested on college students and was shown to be both valid and reliable (Guadagnoli & Mor, 1989). The internal consistency of the scale for the present sample ranged from acceptable to excellent. For the NE subscale, clients' reports of their own emotions were 0.77 and 0.96, for within-client and between-clients, respectively; therapists' reports of their own emotions were 0.68 and 0.95, for within-therapist and between-therapists, respectively; therapists' assessments of their clients' emotions were 0.75 and 0.94, for within-therapist and between-therapists, respectively. For the Positive Emotion subscale, clients' reports of their own emotions were 0.73 and 0.98, for within-client and between-clients, respectively; therapists' reports of their own emotions were 0.82, and 0.97, for within-therapist and between-therapists, respectively; therapists' assessments of their clients' emotions were 0.80 and 0.96, for within-therapist and between-therapists, respectively.

Procedure

The procedures were part of the routine battery in the clinic. Clients were asked to sign consent forms and were told that they could choose to terminate their participation in the study at any time without jeopardizing treatment. Clients were also told that their data would not be shown to their therapist, and their anonymity would be preserved.

The clients completed the HSCL before each therapy session. Both the clients and therapists completed the POMS immediately after each therapy session. The therapists completed the POMS twice; they initially reported their own emotions and subsequently reported how they perceived their clients' emotions during the session. All research materials were collected after securing the approval of the authors' university ethics committee.

Data Analysis Strategy

The dataset had a hierarchical structure, with session ratings nested within clients and clients nested within therapists. Therefore, we used a multilevel model (MLM; Raudenbush & Bryk, 2002), with sessions at Level 1 and clients at Level 2. When we attempted to estimate three-level models (i.e., taking into account therapist effects), the models did not converge. This is likely a result of the average low number of clients treated by the same therapists in our sample (in which most therapists treated only one client); this limits the extent to which we could examine therapist effects.[3] Thus, we retained the two-level model.

To test our first three hypotheses, we used an adaptation of the T&B model (West & Kenny, 2011). In these analyses, the therapists' assessments of their clients' emotions constituted the outcome, which in turn was predicted by (a) the clients' reports of their own emotions and (b) the therapists' reports of their clients' emotions. Thus, the first slope coefficient represented the tracking accuracy; that is, the extent to which the therapists' assessments were temporally accurate regarding their clients' emotions; the second slope coefficient represented the assumed similarity; that is, the extent to which the therapists' assessments regarding their clients' emotions were linked to their own emotions.

In addition, we person-mean-centered the outcome (i.e., the therapists' assessments) and the predictors (i.e., the clients' and therapists' emotions) on the clients' person-mean emotions across all sessions, which enabled us to (a) remove broad individual differences when examining within-person fluctuations and, more importantly, (b) treat the intercept estimate as representing the *directional discrepancy*—that is, the extent to which the therapists,

[2] The decision to aggregate the sub-scales was supported by an examination of between-client and within-client correlations between the subscales of the POMS, which showed that the associations within valences were positive and strong both at the between client (0.62–0.77) and within client (0.39–0.58) levels.

[3] We reran all the analyses with one client per therapist (the one who provided the most session-by-session reports; $N_{clients} = 62$). The pattern of results was almost identical to the ones obtained for the entire sample of clients. These results are available upon request.

on average, overestimate (in cases of positive intercepts) or underestimate (negative intercepts) their clients' emotions.

We ran a multivariate MLM, with NE and PE as outcomes. This approach enabled us to obtain fixed effects for each emotion and the cross-valence covariance between emotions. To do so, we used a dummy variable to denote each outcome. For each outcome, the model included (a) an intercept which represented the directional discrepancy, (b) a slope which represented the tracking accuracy, and (c) a slope which represented the assumed similarity. All estimates were considered to be random, which enabled us to examine their variances and covariances.

In this multivariate multilevel equation, the PE and NE were combined into a single outcome variable, termed therapist's assessment, $_{csj}$, where j indexes the outcome measure. We created two indicator variables (Neg_j and Pos_j), where Neg_j equaled 1 for NE and 0 for PE, and Pos_j equaled 0 for NE and 1 for PE. The equation is as follows:

$$\begin{aligned}\text{Therapist's assessment}_{jsc} = \text{Neg}_j \times [(\gamma_{10} + u_{1c}) + (\gamma_{30} + u_{3c}) \\ \times \text{Client's NE}_{sc} + (\gamma_{50} + u_{5c}) \\ \times \text{Therapist's NE}_{sc} + e_{1sc}] \\ \text{Pos}_j \times [(\gamma_{20} + u_{2c}) + (\gamma_{40} + u_{4c}) \\ \times \text{Client's PE}_{sc} + (\gamma_{60} + u_{6c}) \\ \times \text{Therapist's PE}_{sc} + e_{2sc}],\end{aligned}$$

where the therapist's assessment ($j = 1$ for NE, $j = 2$ for PE, and Pos and Neg as dummy coded variables) in session s with client c were predicted by the following: the sample's average (i.e., fixed) directional discrepancy (i.e., the intercepts; γ_{10} and γ_{20}); the client's NE or PE in this particular session (i.e., the slopes; γ_{30} and γ_{40}); the therapist's NE or PE in this particular session (i.e., the slopes; γ_{50} and γ_{60}); the deviations of this particular therapeutic dyad from the average intercepts and slopes (i.e., the random effects for the directional-discrepancy intercepts u_{1c}, u_{2c}; the random effects for the tracking accuracy slopes: u_{3c}, u_{4c}; the random effects for the assumed similarity slopes: u_{5c}, u_{6c}); and the Level 1 residual terms, which quantify the session's deviation from these effects (i.e., random effects at Level 1; e_{1sc} and e_{2sc}). A first-order autoregressive structure was estimated for the Level 1 random effects, and they were allowed to correlate with each other.

Our first hypothesis concerning therapists' tracking accuracy was tested by the significance tests of the tracking accuracy slope coefficient. Our second hypothesis concerning therapists' over- or underestimation was tested by significance tests of the directional discrepancy intercept. Our third hypothesis concerning the mediation of therapists' accuracy through their own emotions was tested using Bauer, Preacher, and Gil's (2006) method for testing mediation effects in a multilevel model in which the predictor, mediator, and outcome reside at Level 1. Specifically, we ran a multivariate multilevel model which assessed two models simultaneously, the first estimating the association between clients' emotions and therapists' emotions (see path a in Figure 1), and the second estimating the association between therapists' emotions and therapists' assessments (see path b in Figure 1), while controlling for the direct association (i.e., cognitive accuracy) between clients' emotions and therapists' assessments (see path c′ in Figure 1). This method provided estimates for testing the indirect effect (a × b, emotional accuracy), based on both the normal distribution p value and the confidence interval using the Monte Carlo simulation method.

Standardized effect sizes for the directional discrepancies were calculated by dividing the unstandardized estimates by the pooled SDs of the clients' emotions and the therapists' assessments of these emotions and may thus be regarded as an approximation of Cohen's d effect size statistic (see Nezlek, 2012). For the *tracking accuracies* and *assumed similarities*, standardized effect sizes were calculated by standardizing the raw variables and rerunning the models, and may thus be regarded as an approximation of standardized betas (see Baldwin, Imel, Braithwaite, & Atkins, 2014).

To test our fourth hypothesis concerning the association between therapists' EA and changes in clients' symptoms, we used Polynomial Regression with Response Surface Analysis (PRRSA; Edwards & Parry, 1993). This method is well-suited for testing the level of congruence (agreement or accuracy) between two variables as *predictive* of an outcome variable. Note that it overcomes several limitations of traditional methods for testing indices of accuracy (e.g., absolute differences) as predictors (for a review, see Edwards & Parry, 1993). Several psychotherapy research studies have recently recommended PRRSA for this purpose (e.g., Marmarosh & Kivlighan, 2012; Rubel et al., 2018; Zilcha-Mano, Snyder, & Silberschatz, 2016).

In the current study, we estimated a multilevel PRRSA model in which the outcome (i.e., HSCL scores from session $s + 1$) was predicted by the following five variables: (1) the client's emotion in session s; (2) the therapist's assessment of their client's emotion in session s; (3) a first quadratic term that was formed by squaring the client's emotion; (4) a cross-product term that was formed by multiplying the client's emotion by the therapist's assessment; and (5) a second quadratic term that was formed by squaring the therapist's assessment. Before constructing the quadratic and cross-product terms (c, d, & e), the clients' and therapists' reports of their emotions were person-mean centered (Atzil-Slonim et al., 2018; Kivlighan, Li & Gillis., 2015; Shanock, Baran, Gentry, Pattison, & Heggestad, 2010). We also included the outcome level at session s, which allowed us to interpret the outcome as a change score. In addition, to control for therapists' emotions in the same session, we included their reported NE and PE as two additional covariates. Finally, to control for the shared variance between NE and PE, we entered their respective parameters (i.e., two sets of the five parameters described above) into the same model.

The mixed-level equation used to estimate this model appears below; the intercepts, main effects of the client's and therapist's reports, and lagged outcome (symptom levels) were considered to be random at Level 2.[4]

[4] Estimating the cross-product and quadratic terms as random at Level 2 did not improve the model fit.

$$
\begin{aligned}
\mathrm{HSCL}_{(s+1)c} = {} & (\gamma_{00} + u_{0c}) + (\gamma_{10} + u_{1c}) \\
& \times \text{Client's NE}_{sc} + (\gamma_{20} + u_{2c}) \\
& \times \text{Therapist's Assessment of Client's NE}_{sc} + (\gamma_{30}) \\
& \times \text{Client's NE}^2_{sc} + (\gamma_{40}) \times \text{Client's NE} \\
& \times \text{Therapist's Assessment of Client's NE}_{sc} + (\gamma_{50}) \\
& \times \text{Therapist's Assessment of Client's NE}^2_{sc} + (\gamma_{60} + u_{6c}) \\
& \times \text{Client's PE}_{sc} + (\gamma_{70} + u_{7c}) \\
& \times \text{Therapist's Assessment of Client's PE}_{sc} + (\gamma_{80}) \\
& \times \text{Client's PE}^2_{sc} + (\gamma_{90}) \times \text{Client's PE} \\
& \times \text{Therapist's Assessment of Client's PE}_{sc} + (\gamma_{100}) \\
& \times \text{Therapist's Assessment of Client's PE}^2_{sc} + (\gamma_{110} + u_{11c}) \\
& \times \mathrm{HSCL}_{sc} + (\gamma_{120}) \times \text{Therapist's NE}_{sc} + (\gamma_{130}) \\
& \times \text{Therapist's PE}_{sc} + e_{sc}
\end{aligned}
$$

We used the fixed coefficients from the MLM analysis to calculate test values for the four parameters of the positive (or negative) emotion response surfaces (Edwards & Parry, 1993; Shanock et al., 2010), as follows: (1) the linear slope of the line of accuracy (when client's emotions = therapist's assessment of client's emotions; a_1) (2) the curvature along the line of accuracy (a_2) (3) the linear slope of the line of inaccuracy (client's emotions = −[therapist's assessment of client's emotions]; a_3); and (4) the curvature along the line of inaccuracy (a_4).

The a_4 parameter was used to test Hypothesis 4 that inaccuracy (i.e., points farther away from the line of accuracy) would be associated with an increase in clients' symptoms in the following session. Standardized effect sizes were calculated by standardizing the raw variables and rerunning the PRRSA models.

Results

Table 1 reports the means, standard deviations, and a matrix of correlation for the variables. The results of the fixed effects of the T&B model testing Hypotheses 1 through 3 are presented in Table 2. First, as shown in the table and consistent with Hypothesis 1, the parameters estimating tracking accuracy for both NE and PE were significant, indicating that therapists accurately tracked both their clients' PE and NE, although the tracking accuracy was stronger for NE (b = 0.069, *SE* = 0.033, *p* = .038). Second, consistent with Hypothesis 2, the parameters estimating directional discrepancy were significant for both NE and PE, though in opposite directions: therapists tended to overestimate their clients' NE and underestimate their clients' PE.

To rule out the possibility that the positive tracking accuracies or directional discrepancies reflected a positive colinear trend of time (i.e., that both clients' and therapists' emotional ratings increased/decreased over time, which may account for the positive tracking accuracy, directional discrepancy and assumed similarity slopes; see Bolger & Laurenceau, 2013), we repeated the model with time (centered around the middle session) as both a covariate and a moderator of the T&B parameters. In this model, the T&B parameters remained significant and very similar to the ones obtained by the original model for both NE (directional discrepancy: est. = 0.61, *SE* = 0.04; tracking accuracy: est. = 0.28, *SE* = 0.03; assumed similarity: est. = 0.56, *SE* = 0.03, all *p*s < .001) and PE (directional discrepancy: est. = −0.47, *SE* = 0.04; tracking accuracy: est. = 0.20, *SE* = 0.02; assumed similarity: est. = 0.57, *SE* = 0.02, all *p*s < .001). In addition, only the tracking accuracy for NE was moderated by time (est. = −0.004, *SE* = 0.002, *p* = .032), indicating a small decrease in therapists' EA for NE over time.

Third, consistent with Hypothesis 3, clients' emotions and therapists' emotions were positively associated with each other (for NE: b = 0.237, *SE* = 0.026, *p* < .001; for PE: b = 0.166, *SE* = 0.026, *p* < .001). Moreover, also consistent with Hypothesis 3, the parameters estimating the association between therapists' emotions and therapists' assessments of clients' emotions for both NE and PE were significant, indicating that therapists assumed a similarity between their own emotions and their clients' emotions. Finally, again consistent with Hypothesis 3, therapists' own emotions partially mediated the association between clients' emotions and therapists' assessments of NE (indirect effect estimate: 0.133, *SE* = 0.018, *p* < .001, Monte-Carlo CI [0.098, 0.169], 29.4% of total effect) and PE (indirect effect estimate: 0.091, *SE* = 0.015, *p* < .001, Monte-Carlo CI [0.062, 0.122], 27.7% of total effect).[5]

To test Hypothesis 4, we computed a PRRSA model; the results are presented in Table 3. A graphic representation of the response surfaces is presented in Figure 2; Panel A for NE and Panel B for PE; the vertical axes in both panels represent changes in clients' symptoms level (i.e., HSCL) in the next session. The right horizontal axes represent the clients' levels of emotions (NE or PE), and the left horizontal axes represent the therapists' assessments of these emotions. The solid lines represent the lines of accuracy (along which clients' reported emotions and therapists' assessments of these emotion rise jointly). It depicts the levels of clients' symptoms when therapists are accurate in their assessments along a continuum ranging from low levels of emotions (the closest corner) to high levels of emotions (the farthest corner). As can be seen in Table 3, the a_1 parameter testing the linear trend along this line was significant for NE: clients reported higher symptom levels following sessions in which therapists accurately assessed their clients' NE as high than in sessions in which therapists accurately assessed their clients' NE as low. The a_1 parameter was not significant for PE. In addition, the a_2 parameters testing the curvature along the lines of accuracy were not significant for either NE or PE.

The dashed lines in Figure 2 represent the lines of inaccuracy (along which clients' reported emotions rise when therapists' assessments fall). It depicts the levels of clients' symptoms when therapists are inaccurate in their assessments along a continuum ranging from underestimation (the right corner) to overestimation (the left corner). As can be seen in Table 3, the a_3 parameter testing the linear trend along this line was not significant for either NE or PE: clients' reported symptoms levels did not differ following sessions in which their therapists under- or overestimated their emotions. More importantly, and consistent with Hypothesis 4, the a_4 parameter testing the curvature along this line was significant for PE, indicating that moving from the center of this line toward

[5] The effects of the tracking accuracy in models without controlling for the assumed similarity (i.e., the total effects) were as follows: Estimate = 0.388(0.056), *p* < .001; Effect size = .263; for NE, Estimate = 0.277(0.023), *p* < .001, Effect size = .181.

Table 1
Means, Standard Deviations, and Correlation Matrix of the Variables

Variable	1	2	3	4	5
1. Clients' NE	—	−.45***	.27***	−.18***	.37***
2. Clients' PE	−.51***	—	−.25***	.21***	−.21***
3. Therapists' assessment of NE	.41***	−.30**	—	−.31***	.19***
4. Therapists' assessment of PE	−.28**	.23*	−.35***	—	−.12***
5. Clients' HSCL	.64***	−.33**	.22*	−.19	—
M (*SD*)	1.76 (.73)	3.31 (.91)	2.47 (.74)	2.67 (.64)	1.78 (.39)

Note. Between-level correlations are presented below the diagonal and were calculated by averaging the reports over the entire treatment period for each client and therapist. Within-person correlations are presented above the diagonal and were calculated using person-mean centered variables. NE = negative emotions; PE = positive emotions; HSCL = Hopkins Symptom Checklist.
* $p < .05$. ** $p < .01$. *** $p < .001$.

either the right corner (when therapists underestimated their clients' PE), or the left corner (when therapists overestimated their clients' PE) were associated with higher reported symptoms in the following session. This parameter was not significant for NE.

Discussion

Empathy has been extensively investigated in psychotherapy research. However, most studies have used subjective measures of this construct. In contrast, the present work examined the objective phenomenon of therapists' empathic accuracy by assessing therapists' actual ability to accurately infer their clients' states of mind as they fluctuate from session to session over the course of treatment. In doing so, we differentiated between the two components of EA (tracking accuracy and directional discrepancy) and between the accuracies related to positive versus NE. Furthermore, by examining the relative contribution of cognitive and emotional processes, we heeded the repeated calls (e.g., Elliott et al., 2011; Watson, 2016) to differentiate between the emotional and cognitive processes that may lead to therapists' EA. Additionally, we examined the association between EA and treatment outcomes.

To examine our first three hypotheses, we used West and Kenny's (2011) T&B model. The results fully supported our first hypothesis - that therapists would accurately track their clients' negative (and, to a lesser extent, positive) emotions, as well as our second hypothesis - that therapists would tend to overestimate their clients' NE and underestimate their clients' PE. These findings are consistent with accumulating literature on empathic accuracy within psychotherapy (e.g., Barone et al., 2005; Kwon & Jo, 2012).

Therapists' tracking accuracy for NE was significantly higher than their accuracy for PE. This finding is consistent with empathic accuracy studies conducted in other contexts (e.g., close relationships), which have also reported greater accuracy for negative than PE (e.g., Howland & Rafaeli, 2010). This difference may reflect the general primacy of negative stimuli over positive stimuli

Table 2
Fixed and Random Estimates of the Truth and Bias Model

Effect	Fixed effects			
	Estimate (*SE*)	*t* (*df*)	*p*	Effect size
Negative emotions (NE)				
Directional discrepancy (intercept)	.605 (.036)	17.02 (80.6)	<.001	.877
Tracking accuracy	.272 (.030)	8.95 (61.6)	<.001	.185
Assumed similarity	.561 (.027)	20.42 (62.5)	<.001	.553
Positive emotions (PE)				
Directional discrepancy (intercept)	−.472 (.036)	−13.21 (80.5)	<.001	.610
Tracking accuracy	.203 (.020)	9.98 (74.8)	<.001	.133
Assumed similarity	.568 (.022)	25.30 (82)	<.001	.643

	Random effects					
	1	2	3	4	5	6
1. NE: Directional discrepancy	.098 (.018)***	.024 (.012)*	−.011 (.010)	−.031 (.031)*	.002 (.007)	−.002 (.008)
2. NE: Tracking accuracy	.428	.032 (.012)**	.003 (.008)	−.003 (.011)	.007 (.005)	.008 (.006)
3. NE: Assumed similarity	−.237	.114	.021 (.010)*	.005 (.010)	−.004 (.006)	.013* (.006)
4. PE: Directional discrepancy	−.317	−.046	.105	.098 (.018)***	.006 (.008)	.015 (.008)
5. PE: Tracking accuracy	.056	.363	−.235	.178	.011 (.005)**	.001 (.005)
6. PE: Assumed similarity	−.043	.314	.621	.322	.036	.021 (.007)***

Note. The bottom panel of the table provides the random effects estimates. Covariances (with standard errors in parentheses) are provided above the diagonal; variances (with standard errors in parentheses) are provided on the diagonal; standardized correlations are provided below the diagonal.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 3
Response Surfaces for Client Emotions and Therapists' Assessment of Clients' Emotions as Predictors of Next-Session Level Outcomes

Effect	Negative emotions			Positive emotions		
	Estimate (*SE*)	*p*	Standard estimate	Estimate (*SE*)	*p*	Standard estimate
Fixed effects						
Clients' emotions	.06 (.03)	.023	.05	.03 (.02)	.088	.04
Therapists' assessments	.01 (.02)	.529	.01	−.01 (.02)	.700	−.01
Clients' emotions^2	−.02 (.02)	.297	−.01	.01 (.02)	.524	.02
Clients' Emotions × Therapists' assessments	.00 (.04)	.999	−.01	−.08 (.04)	.025	−.06
Therapists' assessments^2	−.03 (.02)	.258	−.02	.05 (.02)	.039	.04
Response surface parameters						
a_1	.07 (.03)	.019	.06	.02 (.03)	.361	.04
a_2	−.05 (.03)	.089	−.04	−.02 (.03)	.599	−.001
a_3	.05 (.03)	.160	.04	.04 (.03)	.189	.05
a_4	−.05 (.06)	.443	−.02	.14 (.06)	.018	.11

Note. The model also included the intercept, the therapists' own reported negative emotions and positive emotions, and the clients' reported symptoms during the previous week. Standardized estimates were obtained by standardizing the raw scores and rerunning the model. a_1 = the linear slope of the line of agreement/accuracy; a_2 = the curvature along the line of agreement/accuracy; a_3 = the linear slope of the line of disagreement/inaccuracy; a_4 = the curvature along the line of disagreement/inaccuracy.

(Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Gottman, 1994); if the signal value of negativity is stronger, empathic inferences concerning it may be based on more abundant information. Indeed, clients typically enter therapy with some distress, and this distress is likely to take center stage. Moreover, therapists may be more attuned to NE because of the higher potential cost of missing these emotions (compared with missing PE). For example, failing to identify a client's anger directed at the therapist may lead to a substantial therapeutic rupture, which in turn may lead to negative consequences or an early dropout from therapy. In comparison, the failure to identify a client's contentment or excitement is likely to have lower costs.

The marked difference between accuracy regarding NE and PE was also evident in the directional discrepancy findings. As expected, therapists underestimated PE and overestimated NE. These findings appear to cohere with those reported in Atzil-Slonim et al. (2018), who found that therapists experienced less intense PE than their clients, but did not differ from their clients in the intensity of NE. It is also consistent with recent reviews (Fletcher & Kerr, 2010) and empirical work (e.g., Overall et al., 2012; Sened et al.,

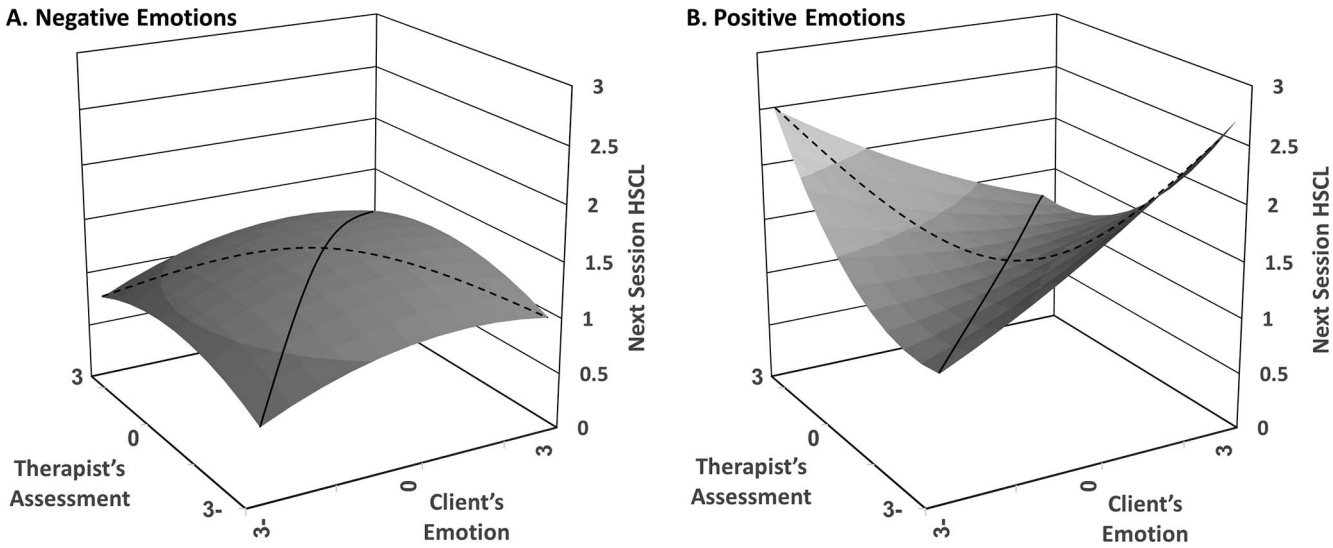

Figure 2. Polynomial regression with response surface analysis models predicting clients' next session symptoms by clients' emotions and therapists' assessments of their clients' emotions. The dashed lines represent the lines of inaccuracy. HSCL = Hopkins Symptom Checklist.

2017) showing that in close relationships, partners tend to adopt a somewhat pessimistic approach when they assess each other's mental states because this approach may be associated with relational benefits.

Our third hypothesis explored the relative importance of two processes to accuracy: one cognitive and the other emotional (see Figure 1). The findings offered support for both processes. Specifically, therapists' PE and NE were temporally similar to their clients' same-valence emotions, and partially mediated their tracking accuracy for both types of emotion; in other words, therapists used emotional accuracy—they (correctly) assumed that their experience was similar to their clients, and this accounted for part of the correct inference. At the same time, their cognitive accuracy (i.e., the accuracy of therapists' assessments above and beyond the mediation through real and assumed similarity) also remained significant.

The co-occurrence of two processes to accuracy is consistent with affective neuroscience studies (Shamay-Tsoory, 2011; Zaki & Ochsner, 2012) that speak to the existence of both cognitive empathy (considered to be the product of a mental state attribution system) and emotional empathy (considered to be the product of an experience sharing system). Interestingly, the co-occurrence of these processes has also been documented in studies assessing therapists' subjective reports: specifically, therapists who were asked to identify the means they used to infer their clients' experiences noted both cognitive means (e.g., attending to the client's story and body language) and emotional means (e.g., noting their own bodily reaction or feelings as they unfold throughout the session; Greenberg & Rushanski-Rosenberg, 2002).

Notably, the cognitive EA (involving the unmediated process) was more than twice as strong as the emotional EA (involving real and assumed similarity). Outside of psychotherapy, studies that have compared the relative strength of cognitive and emotional processes to accuracy (e.g., between romantic partners; Sened et al., 2017; Wilhelm & Perrez, 2004) have tended to indicate a stronger role for the emotional process. Our findings may differ because of the nature of the relationship under study. The inherent closeness between romantic partners may lead to greater use of the emotional EA process. In contrast, even though therapists appear to experience similar emotions as their clients, they are likely to maintain some level of objectivity and not rely solely on these emotions when inferring their clients' emotions. Indeed, there is evidence that therapists are more effective than laypeople in regulating their NE (Pletzer, Sanchez, & Scheibe, 2015). Thus, they may experience the client's emotion during the session, but recover by the end of the session and still use this emotional information to accurately assess their client's emotions. Future studies may examine this possibility by including more fine-grained assessments of both clients' and therapists' emotions (e.g., by using external observers' ratings of the dyad's emotions moment by moment).

Another possible explanation for the difference between our findings and those in the romantic relationships literature is that in therapy, clients may be asked for (and may indeed provide) more explicit verbal content about their in-session feelings than is customary in other contexts (including romantic relationships). Additionally, when therapists assess their clients' feelings (perhaps through emotional empathy) they often verify it using explicit questions which may reinforce their cognitive empathy.

To examine our fourth hypothesis (i.e., the association between therapists' EA and symptom change), we used PRRSA (Edwards & Parry, 1993), a statistical model particularly well-suited for examining the predictive value of EA. In partial support of our hypothesis, the results indicated that higher therapist EA regarding clients' PE (but not regarding clients' NE) was associated with lower symptom reports in the following session. These findings are in line with Kwon and Jo (2012), who identified an association between EA levels and clients' perception of counseling outcome, as well as with Duan and Kivlighan (2002), who reported an association between both cognitive and emotional processes to EA on the one hand, and clients' ratings of session depth on the other. The current study complements these earlier ones by using full session-by-session data.

It is interesting to speculate why empathic accuracy regarding PE—but not regarding NE—was associated with symptomatic improvement from session to session. We argued in the preceding text that therapists' greater tracking accuracy of NE (compared with PE), alongside their (average) tendency to underestimate PE, reflected a greater signal value for negativity. However, these factors could also be seen as evidence of therapists' general neglect of PE—a phenomenon that has received recent theoretical attention in psychotherapy research and clinical practice (Stalikas, Fitzpatrick, Mistkidou, Boutri, & Seryianni, 2015). The finding that inaccuracy in PE was associated with higher symptoms lends credence to the possibility that this neglect comes at a price.

Finally, we found a strong positive association between therapists' tracking accuracies for PE and NE. This finding may suggest that some clients are more readable than others or that some therapists are better attuned to changes in emotion, regardless of valence. This finding, regarding which we had not a priori hypotheses, warrants further exploration.

Limitations, Future Directions, and Clinical Implications

One limitation of this study is its use of aggregated total scores for PE and NE. There is evidence that the distinction between EA for specific NE (e.g., anger vs. sadness) may be meaningful (e.g., Schoebi & Randall, 2015). Furthermore, by using the aggregated PE and NE, we may erroneously deem some instances of inaccuracy as indicative of accuracy. For example, if a client experiences elevated sadness whereas his therapist views it as fear, the therapist may be seen, incorrectly, as empathically accurate. This limitation notwithstanding, we opted to use the aggregated scores given the finding of medium-to-large associations among the specific emotions within each valence (average = .581; compared with an average of −.343 for cross-valence correlations). Future studies should examine different potential patterns of EA for specific emotions.

Another limitation inherent to the way we operationalized EA was that it relied entirely on numerical reports of emotions, thus eschewing other important parts of experience (e.g., the specific content of clients' thoughts). Moreover, our operationalization led us to study emotions at a relatively low time resolution (once each session, typically weekly), even though emotions often fluctuate at a much higher time resolution (Butler, 2015). The reliance on client and therapist self-reports of emotions could have biased our results. For example, clients or therapists might have been de-

fended against NE or were unwilling to admit overly NE because of social desirability concerns. The act of completing measures of emotions after each session may have also increased therapist and client self-awareness of emotions and influenced therapy process and outcome. Finally, when therapists are asked to rate their own emotions and assess their clients' emotions, there may be some kind of anchoring effect that does not reflect their real ability to be empathically accurate with their clients. Future studies may assess EA during sessions using measures other than self-reports and may benefit from examining associations among various operationalizations of EA—as well as between such operationalizations and therapy outcomes.

The therapists in this study were trainees in a program that emphasizes a psychodynamic model of treatment, in which the key features include a focus on affect and the experience and expression of emotions. Both of these factors (the therapists' experience level and the emphasis on emotions in their training) may limit the generalizability of the findings. Although we consider EA to be a pan-theoretical component emphasized by most psychotherapy orientations, future studies are required to explore this (and determine its consequences) with therapists implementing other therapeutic orientations (and at other training levels).

Because most therapists in our sample treated only one client each, we could not estimate therapist effects. Future studies with larger numbers of clients per therapist are necessary to examine specific therapist characteristics that may moderate the therapists' ability to be empathically accurate, such as their flexible versus rigid repertoire of emotional style (Holmqvist, 2001). Similarly, recent reviews have indicated that target individuals vary in the overall "readability" of their feelings (Hall et al., 2016); thus, future studies could explore the specific characteristics that make certain clients easier to read. Future studies may also benefit from identifying the conditions in which EA leads or fails to lead to positive outcomes in psychotherapy. For example, it would be interesting to examine whether in situations of alliance rupture, therapists' EA toward clients' emotions has a stronger impact on session outcome. Outside of psychotherapy research, this moderation has been demonstrated. For example, in a study of romantic couples, Lazarus, Bar-Kalifa, and Rafaeli (2017) reported that EA was associated more strongly to relationship outcomes on days in which conflict occurred.

Several clinical implications may be drawn from this work. First, our results advance the idea that both cognitive and emotional empathy contribute to therapists' ability to correctly assess their clients' emotions. Ideally, therapists should use both cognitive and emotional means to arrive at EA. However, it is likely that some therapists have different levels of these abilities. Our results may imply that novice therapists may be taught to read affect and be more attentive to emotional cues by using cognitive or emotional paths.

Second, our results suggest that therapists who tend to be accurate in assessing their clients' painful emotional experiences are likely to be the same therapists who tend to be accurate in assessing their clients' positive experiences. Alternatively, it is possible that clients whose PE are easily detected also tend to be the same clients whose NE are more easily detected than other clients. This finding calls for future studies with samples that include more clients per therapist to examine whether some therapists tend to be more sensitive than others or whether some clients tend to be more easily detected, and to identify therapist and client characteristics associated with EA.

Finally, our results indicate that overall, PE tends to be neglected by therapists (as reflected by the higher accuracy for NE than PE and the significant overestimation of NE and underestimation of PE). Moreover, this neglect is associated with treatment outcome. By underestimating the importance of PE, therapists may miss opportunities to use these emotions as a springboard to advance positive change. The current findings thus enhance our understanding of specific characteristics associated with more (or less) EA and may help therapists to be more attuned to their clients' emotions.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5)*. Washington, DC: Author.

Atzil-Slonim, D., Bar-Kalifa, E., Fisher, H., Peri, T., Lutz, W., Rubel, J., & Rafaeli, E. (2018). Emotional congruence between clients and therapists and its effect on treatment outcome. *Journal of Counseling Psychology, 65,* 51–64. http://dx.doi.org/10.1037/cou0000250

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

Baldwin, S. A., Imel, Z. E., Braithwaite, S. R., & Atkins, D. C. (2014). Analyzing multiple outcomes in clinical research using multivariate multilevel models. *Journal of Consulting and Clinical Psychology, 82,* 920–930. http://dx.doi.org/10.1037/a0035628

Barone, D. F., Hutchings, P. S., Kimmel, H. J., Traub, H. L., Cooper, J. T., & Marshall, C. M. (2005). Increasing empathic accuracy through practice and feedback in a clinical interviewing course. *Journal of Social and Clinical Psychology, 24,* 156–171. http://dx.doi.org/10.1521/jscp.24.2.156.62275

Barrett-Lennard, G. T. (1981). The empathy cycle: Refinement of a nuclear concept. *Journal of Consulting and Clinical Psychology, 28,* 91–100.

Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods, 11,* 142–163. http://dx.doi.org/10.1037/1082-989X.11.2.142

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5,* 323–370. http://dx.doi.org/10.1037/1089-2680.5.4.323

Blagys, M. D., & Hilsenroth, M. J. (2000). Distinctive features of short-term psychodynamic-interpersonal psychotherapy: A review of the comparative psychotherapy process literature. *Clinical Psychology: Science and Practice, 7,* 167–188. http://dx.doi.org/10.1093/clipsy.7.2.167

Bohart, A. C., & Greenberg, L. S. (1997). *Empathy reconsidered: New directions in psychotherapy*. Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/10226-000

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

Butler, E. A. (2015). Interpersonal affect dynamics: It takes two (and time) to tango. *Emotion Review, 7,* 336–341. http://dx.doi.org/10.1177/1754073915590622

Chui, H., Hill, C. E., Kline, K., Kuo, P., & Mohr, J. J. (2016). Are you in the mood? Therapist affect and psychotherapy process. *Journal of Counseling Psychology, 63,* 405–418. http://dx.doi.org/10.1037/cou0000155

Cohen, S., Schulz, M. S., Weiss, E., & Waldinger, R. J. (2012). Eye of the beholder: The individual and dyadic contributions of empathic accuracy and perceived empathic effort to relationship satisfaction. *Journal of Family Psychology, 26,* 236–245. http://dx.doi.org/10.1037/a0027488

Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change:

Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin, 32,* 917–929. http://dx.doi.org/10.1177/0146167206287721

Cuff, B. M., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A review of the concept. *Emotion Review, 8,* 144–153. http://dx.doi.org/10.1177/1754073914558466

Decety, J., & Ickes, W. (2011). *The social neuroscience of empathy*. Cambridge, MA: MIT Press.

Derogatis, L. R. (1975). *Brief symptom inventory*. Baltimore, 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

Duan, C., & Hill, C. E. (1996). The current state of empathy research. *Journal of Counseling Psychology, 43,* 261–274. http://dx.doi.org/10.1037/0022-0167.43.3.261

Duan, C., & Kivlighan, D. M. (2002). Relationships among therapist presession mood, therapist empathy, and session evaluation. *Psychotherapy Research, 12,* 23–37. http://dx.doi.org/10.1080/713869615

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.

Elliott, R., Bohart, A. C., Watson, J. C., & Greenberg, L. S. (2011). Empathy. *Psychotherapy: Theory, Research, & Practice, 48,* 43–49. http://dx.doi.org/10.1037/a0022187

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

Fosha, D. (2001). The dyadic regulation of affect. *Journal of Clinical Psychology, 57,* 227–242. http://dx.doi.org/10.1002/1097-4679(200102)57:2<227::AID-JCLP8>3.0.CO;2-1

Gottman, J. (1994). Why marriages fail. *The Family Therapy Networker, 18,* 40–48.

Greenberg, L. S. (2010). The clinical application of emotions in psychotherapy. In M. Lewis, L. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (pp. 88–101). New York, NY: Guilford Press.

Greenberg, L. S., & Elliott, R. (1997). Varieties of empathic responding. In A. Bohart & L. S. Greenberg (Eds.), *Empathy reconsidered* (pp. 167–186). Washington, DC: APA Books.

Greenberg, L. S., & Rushanski-Rosenberg, R. (2002). Therapist's experience of empathy. In J. C. Watson, R. N. Goldman, & M. S. Warner (Eds.), *Client-centered and experiential psychotherapy in the 21st century: Advances in theory, research, and practice* (pp. 168–181). Ross-on Wye, UK: PCCS Books.

Guadagnoli, E., & Mor, V. (1989). Measuring cancer patients' affect: Revision and psychometric properties of the Profile of Mood States (POMS). *Psychological Assessment, 1,* 150–154. http://dx.doi.org/10.1037/1040-3590.1.2.150

Hall, J. A., Mast, M. S., & West, T. V. (2016). *The social psychology of perceiving others accurately*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781316181959

Hayes, A. M., Ready, C. B., & Yasinski, C. (2014). Application of exposure and emotional processing theory to depression: Exposure-based cognitive therapy. In N. C. Thoma & D. McKay (Eds.), *Working with emotion in cognitive-behavioral therapy: Techniques for clinical practice*. New York, NY: Guilford Press.

Hayes, J. A., Gelso, C. J., & Hummel, A. M. (2011). Managing countertransference. *Psychotherapy: Theory, Research, & Practice, 48,* 88–97. http://dx.doi.org/10.1037/a0022182

Hodges, S. D., & Wegner, D. M. (1997). The mental control of empathic accuracy. In W. Ickes (Ed.), *Empathic accuracy* (pp. 311–339). New York, NY: Guilford Press.

Hofmann, S. G., & Hayes, S. C. (2015). *Emotion in therapy: From science to practice*. New York, NY: Guilford Press Publications.

Holmqvist, R. (2001). Patterns of consistency and deviation in therapists' countertransference feelings. *The Journal of Psychotherapy Practice and Research, 10,* 104–116.

Holmqvist, R., Hansjons-Gustafsson, U., & Gustafsson, J. (2002). Patients' relationship episodes and therapists' feelings. *Psychology and Psychotherapy: Theory, Research and Practice, 75,* 393–409. http://dx.doi.org/10.1348/147608302321151907

Howland, M., & Rafaeli, E. (2010). Bringing everyday mind reading into everyday life: Assessing empathic accuracy with daily diary data. *Journal of Personality, 78,* 1437–1468. http://dx.doi.org/10.1111/j.1467-6494.2010.00657.x

Ickes, W. J. (2003). *Everyday mind reading: Understanding what other people think and feel*. Amherst, NY: Prometheus Books.

Ickes, W., & Hodges, S. D. (2013). Empathic accuracy in close relationships. In J. A. Simpson & L. Campbell (Eds.), *The Oxford handbook of close relationships* (pp. 348–373). New York, NY: Oxford University Press.

Kenny, D. A., & Acitelli, L. K. (2001). Accuracy and bias in the perception of the partner in a close relationship. *Journal of Personality and Social Psychology, 80,* 439–448. http://dx.doi.org/10.1037/0022-3514.80.3.439

Keown, L. J., & Woodward, L. J. (2002). Early parent–child relations and family functioning of preschool boys with pervasive hyperactivity. *Journal of Abnormal Child Psychology, 30,* 541–553. http://dx.doi.org/10.1023/A:1020803412247

Kivlighan, D. M., Jr., Li, X., & Gillis, L. (2015). Do I fit with my group? Within-member and within-group fit with the group in engaged group climate and group members feeling involved and valued. *Group Dynamics: Theory, Research, and Practice, 19,* 106. http://dx.doi.org/10.1037/gdn0000025

Kohut, H. (1977). *The Restoration of the self*. New York, NY: International Universities Press.

Kwon, K. I., & Jo, S. Y. (2012). The relationship among counselor experience level, empathic accuracy, and counseling outcome in the early phase of counseling. *Asia Pacific Education Review, 13,* 771–777. http://dx.doi.org/10.1007/s12564-012-9235-8

Lazarus, G., Bar-Kalifa, E., & Rafaeli, E. (2017). Accurate where it counts: Empathic accuracy on conflict and no-conflict days. *Emotion, 18,* 212–228. http://dx.doi.org/10.1037/emo0000325

LeDoux, J. E., & Hofmann, S. G. (2018). The subjective experience of emotion: A fearful view. *Current Opinion in Behavioral Sciences, 19,* 67–72. http://dx.doi.org/10.1016/j.cobeha.2017.09.011

Lutz, W., Tholen, S., Schürch, E., & Berking, M. J. B. (2006). Die entwicklung, validität und reliabilität von kurzformen gängiger psychometrischer instrumente zur evaluation des therapeutischen fortschrittes 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

Main, A., Walle, E. A., Kho, C., & Halpern, J. (2017). The interpersonal functions of empathy: A relational perspective. *Emotion Review, 9,* 358–366. http://dx.doi.org/10.1177/1754073916669440

Marangoni, C., Garcia, S., Ickes, W., & Teng, G. (1995). Empathic accuracy in a clinically relevant setting. *Journal of Personality and Social Psychology, 68,* 854–869. http://dx.doi.org/10.1037/0022-3514.68.5.854

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

McCullough, L., Kuhn, N., Andrews, S., Kaplan, A., Wolf, J., & Hunsley, C. L. (2003). *Treating affect phobia: A manual for short-term dynamic psychotherapy*. New York, NY: Guilford Press.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1992). *POMS manual—Profile of mood questionnaire*. San Diego, CA: EdITS.

Nezlek, J. (2012). Multilevel modeling analyses of diary-style data. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 357–383). New York, NY: Guilford Press.

Overall, N. C., Fletcher, G. J., & 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

Overall, N. C., Fletcher, G. J., Simpson, J. A., & Fillo, J. (2015). Attachment insecurity, biased perceptions of romantic partners' negative emotions, and hostile relationship behavior. *Journal of Personality and Social Psychology, 108,* 730–749. http://dx.doi.org/10.1037/a0038987

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience, 9,* 148–158. http://dx.doi.org/10.1037/a0039078

Pletzer, J. L., Sanchez, X., & Scheibe, S. (2015). Practicing psychotherapists are more skilled at downregulating negative emotions than other professionals. *Psychotherapy: Theory, Research, & Practice, 52,* 346–350. http://dx.doi.org/10.1037/a0039078

Preston, S. D., & Hofelich, A. J. (2012). The many faces of empathy: Parsing empathic phenomena through a proximate, dynamic-systems view of representing the other in the self. *Emotion Review, 4,* 24–33. http://dx.doi.org/10.1177/1754073911421378

Rafaeli, E., Gadassi, R., Howland, M., Boussi, A., & Lazarus, G. (2017). Seeing bad does good: Relational benefits of accuracy regarding partners' negative moods. *Motivation and Emotion, 41,* 353–369. http://dx.doi.org/10.1007/s11031-017-9614-x

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: SAGE.

Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology, 21,* 95–103. http://dx.doi.org/10.1037/h0045357

Rubel, J., Bar-Kalifa, E., Atzil-Slonim, D., Schmidt, S., & Lutz, W. (2018). Congruence of therapeutic bond perceptions and its relation to treatment outcome: Within- and between-dyad effects. *Journal of Consulting and Clinical Psychology, 86,* 341–353. http://dx.doi.org/10.1037/ccp0000280

Schoebi, D., & Randall, A. K. (2015). Emotional dynamics in intimate relationships. *Emotion Review, 7,* 342–348. http://dx.doi.org/10.1177/1754073915590620

Sened, H., Yovel, I., Bar-Kalifa, E., Gadassi, R., & Rafaeli, E. (2017). Now you have my attention: Empathic accuracy pathways in couples and the role of conflict. *Emotion, 17,* 155–168. http://dx.doi.org/10.1037/emo0000220

Shamay-Tsoory, S. G. (2011). Empathic processing: Its cognitive and affective dimensions and neuroanatomical basis. In J. Decety & W. Ickes (Eds.), *The social neuroscience of empathy* (pp. 215–232). Cambridge, MA: MIT Press.

Shanock, L. R., Baran, B. E., Gentry, W. A., Pattison, S. C., & Heggestad, 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

Shedler, J. (2010). The efficacy of psychodynamic psychotherapy. In R. A. Levy, J. S. Ablon, & H. Kachele (Eds.), *Psychodynamic psychotherapy research: Evidence-based practice and practice-based evidence* (pp. 9–25). New York, NY: Humana Press.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., . . . Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for *DSM–IV* and ICD-10. *The Journal of Clinical Psychiatry, 59*(Suppl. 20), 22–33.

Stalikas, A., Fitzpatrick, M., Mistkidou, P., Boutri, A., & Seryianni, C. (2015). Positive emotions in psychotherapy: Conceptual propositions and research challenges. In O. C. G. Gelo, A. Pritz, & B. Rieken (Eds.), *Psychotherapy research: Foundations, process, and outcome* (pp. 331–349). Vienna, Austria: Springer Vienna. http://dx.doi.org/10.1007/978-3-7091-1382-0_17

Summers, R. F., & Barber, J. P. (2010). *Psychodynamic therapy: A guide to evidence-based practice*. New York, NY: Guilford Press.

Thwaites, R., & Bennett-Levy, J. (2007). Conceptualizing empathy in cognitive behaviour therapy: Making the implicit explicit. *Behavioural and Cognitive Psychotherapy, 35,* 591–612. http://dx.doi.org/10.1017/S1352465807003785

Watson, J. C. (2016). The role of empathy in psychotherapy: Theory, research, and practice. In D. J. Cain, K. Keenan, & S. Rubin (Eds.), *Humanistic psychotherapies: Handbook of research and practice* (pp. 115–145). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/14775-005

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

Westra, H. A., Aviram, A., Connors, L., Kertes, A., & Ahmed, M. (2012). Therapist emotional reactions and client resistance in cognitive behavioral therapy. *Psychotherapy: Theory, Research, & Practice, 49,* 163–172. http://dx.doi.org/10.1037/a0023200

Wilhelm, P., & Perrez, M. (2004). How is my partner feeling in different daily-life settings? Accuracy of spouses' judgements about their partner's feelings at work and at home. *Social Indicators Research, 67,* 183–246. http://dx.doi.org/10.1023/B:SOCI.0000007339.48649.20

Zaki, J., & Ochsner, K. N. (2012). The neuroscience of empathy: Progress, pitfalls and promise. *Nature Neuroscience, 15,* 675–680. http://dx.doi.org/10.1038/nn.3085

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. http://dx.doi.org/10.1080/10503307.2015.1126682

Received July 24, 2017
Revision received June 16, 2018
Accepted August 22, 2018 ■