

# In-Session Reflective Functioning in Psychotherapies for Borderline Personality Disorder: The Emotion Regulatory Role of Reflective Functioning

Yogev Kivity<sup>1</sup>, Kenneth N. Levy<sup>1, 2</sup>, Kristen M. Kelly<sup>1</sup>, and John F. Clarkin<sup>2</sup>

<sup>1</sup> Department of Psychology, Pennsylvania State University

<sup>2</sup> Joan and Sanford I. Weill Medical College of Cornell University

**Objective:** The capacity for understanding mental states (reflective functioning; RF) is considered essential for self-growth, social learning, and emotion regulation. Impaired RF is thought to play a central role in borderline personality disorder (BPD). We examined whether asking patients to consider mental states in-session has a down-regulatory effect on emotional arousal in treatments for BPD. **Method:** Early-, middle- and late-phase videotaped sessions from a randomized-controlled trial of transference-focused psychotherapy (TFP;  $n = 30$ ), dialectical behavior therapy (DBT;  $n = 29$ ), and supportive psychodynamic therapy (SPT;  $n = 29$ ) were segmented to therapist and patient talk-turns. Therapist talk-turns were rated as asking patients to consider mental state (bids for RF) or not. Patient talk-turns were rated for RF and acoustically encoded for arousal. **Results:** Bids were twice as common in TFP compared to DBT and SPT. Across treatments, therapist bids for RF predicted better RF, which, in turn, predicted lower emotional arousal. **Conclusions:** Asking patients to consider mental states has a down-regulatory effect on patients' arousal in psychotherapies for BPD.

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


In psychotherapies for borderline personality disorder (BPD), therapist interventions that ask patients to consider mental states are associated with increased capacity for understanding mental states which, in turn, is associated with reduced emotional arousal. Thus, understanding mental states might have an emotion regulatory role in psychotherapies for BPD.

**Keywords:** borderline personality disorder, transference-focused psychotherapy, dialectical behavior therapy, supportive psychodynamic therapy, reflective functioning

**Supplemental materials:** <https://doi.org/10.1037/ccp0000674.supp>

Yogev Kivity  <https://orcid.org/0000-0002-8293-8910>

Kenneth N. Levy  <https://orcid.org/0000-0002-4670-4631>

Kristen M. Kelly  <https://orcid.org/0000-0001-5967-8367>

Yogev Kivity is now in the Department of Psychology, Bar Ilan University, Israel.

John F. Clarkin receives royalties from a book published by American Psychiatric Publishing. All other authors report no conflicting interests.

Levy receives royalties from books published by Sage Publications, Routledge, and Elsevier publishers.

This research was supported by grants from the National Institute of Mental Health, International Psychoanalytic Association, the American Psychoanalytic Association, the Research and Graduate Studies Office (RGSO) in the College of the Liberal Arts at the Pennsylvania State University and the Kohler Fund of Munich awarded to Kenneth N. Levy, and a grant from the Borderline Personality Disorder Research Foundation (BPDFR) awarded to Otto F. Kernberg and John F. Clarkin. Preparation of this manuscript was also supported by grants from the Fulbright program (PI: Yogev Kivity), The American Psychoanalytic Association (PI: Yogev Kivity), The International Psychoanalytic Association (PI: Yogev Kivity; co-PIs: Kenneth N. Levy and John F. Clarkin) and The American Psychological Association, Division 29 (PI: Yogev Kivity; co-PIs: Kenneth N. Levy and John F. Clarkin). We thank the BPDFR founder and executive officer, Marco Stoffel, and the scientific board for their advice and encouragement. We

thank Jack Barchas and Vita Rabinowitz for their respective institutional support. In addition, we acknowledge the technical assistance of Kenneth L. Critchfield, Pamela E. Foelsch, Simone Hoermann, Maya K. Krischer, and Joel McClough as well as Jill C. Delaney for their help in conducting assessments, and James Hull for organizing and maintaining the data in the early stages of the project. We thank Tracy L. Clouthier, Thomas M. Crow, Benjamin N. Johnson, and Lia K. Rosenstein for coding in-session reflective functioning. We also thank Kevin B. Meehan, Joseph S. Reynoso, and Michal Weber for coding Adult Attachment Interview (AAI) reflective functioning. We thank Megan N. Parker and Jessica K. Hlay for coordinating coding for the study and Aidan H. McMorrow, Sona Hussain, and Macy Celitti for coding therapist bids for reflection. We thank members of the Personality Disorders Institute. We thank the therapists in the study, and Frank Yeomans, Ann Appelbaum, and Barbara Stanley for serving as treatment cell leaders and providing supervision. Finally, we thank the patients for their participation.

The data and study materials are not available as they include personally identifiable information. The data analysis code is available upon request from the first author.

Correspondence concerning this article should be addressed to Yogev Kivity, Department of Psychology, Bar Ilan University, 304 Max & Anna Webb Building, Ramat Gan, 5290002, Israel or Kenneth N. Levy, Department of Psychology, Pennsylvania State University, 362 Bruce V. Moore Building, University Park, PA 16802, United States. Email: [yogev.kivity@biu.ac.il](mailto:yogev.kivity@biu.ac.il) or [klevy@psu.edu](mailto:klevy@psu.edu)



Borderline personality disorder (BPD) is a severe and prevalent disorder characterized by instability in identity, emotion, and interpersonal relationships (Skodol et al., 2002). Several specialized treatments are efficacious in reducing symptoms of BPD, but treatment effects tend to be moderate to small (Cristea et al., 2017). In addition, specialized treatments for BPD do not differ in their efficacy in reducing symptoms (e.g., Cristea et al., 2017) despite the different theorized mechanisms of action underlying these treatments (Clarkin & Levy, 2006). Studying the unique and common processes of change in treatments of BPD may help shed light on the actual mechanisms underlying these treatments and may ultimately help improve their outcome.

### The Role of Reflective Functioning in Treatments of BPD

Reflective functioning (RF) is an empirical operationalization of the concept of mentalizing—the ability to consider self- and other's mental states. RF is one potential mechanism of change in treatments for BPD. The capacity for RF is postulated to be a fundamental and crucial human function. Inferring intentions through RF is hypothesized to be essential for social learning and self-organization and is assumed to be achieved through the processing of interpersonal and emotional experiences (Fonagy, 2001). Understanding mental states makes the world and relationships with others more predictable and meaningful (Fonagy et al., 1998). It allows one to distinguish between appearance and reality, safely explore one's environment, maintain close relationships, benefit from the social world, appropriately trust (or distrust) others, and effectively tolerate and manage distressing mental states and emotions (Fonagy, 2001). In support of the notion of RF as a fundamental human function, a recent comprehensive meta-analysis found that better capacity for RF predicts positive outcomes in a host of life domains (Kivity et al., 2021).

Importantly, numerous studies suggest that the capacity for RF is impaired in individuals with BPD (see Fonagy & Luyten, 2009 for a review). These findings, along with theoretical considerations, led some authors to propose that change in RF is a presumed mechanism of change that is specific to psychodynamic therapies (Fonagy & Bateman, 2006; Levy, Clarkin, et al., 2006) while other authors propose that it is shared across various treatments for BPD (Bateman et al., 2018). To our knowledge, the existing evidence suggests that transference-focused psychotherapy (TFP; Yeomans et al., 2015), a psychodynamic treatment for BPD, appears to be uniquely associated with increases in RF (Fischer-Kern et al., 2015; Levy, Meehan, et al., 2006; Levy et al., 2021) while other approaches do not show such increases (e.g., Möller et al., 2017).

These findings raise important questions regarding the active ingredients (i.e., therapist interventions) that underlie this specific increase in RF in TFP and the way this increase unfolds over time. One recent study found that general conformity to principles of TFP (and not other principles) predicted increased RF in TFP but not in other treatments (Kivity et al., 2019). This finding suggests that some aspects of TFP indeed play a role in promoting increased RF. In light of these findings, an examination of specific ingredients of TFP has the potential to provide insights about the way these changes are brought about. Specifically, Levy, Clarkin, et al. (2006) proposed that one such mechanism is that TFP therapists tend to utilize interventions that ask patients to consider mental states ("bids" or appeals for RF), for example, by asking patients

what they think went in other people's mind when they interacted with them. These bids are hypothesized to increase the likelihood of patients engaging in RF and lead to increased capacity for RF. The increased RF is then hypothesized to promote increased psychological integration, leading to adaptive emotional processing, and ultimately lead to a reduction in BPD symptoms (Levy, Clarkin, et al., 2006). Such dynamic exchanges between therapist and patient unfold moment-by-moment within sessions. Therefore, microanalytic designs that tap into moment-by-moment processes of change are essential for capturing these dynamics.

In a recent study, Möller et al. (2017) utilized such a microanalytic design to examine in-session RF during mentalization-based treatment (MBT) for patients with BPD and comorbid substance abuse disorder ( $n = 5$ ). Individual therapist and patient talk-turns were coded for use of bids and for RF, respectively. The findings showed that therapist use of bids predicted better RF in the talk-turn that followed the bid. This important finding provides preliminary support to the notion that therapist bids for RF may explain increases in RF in treatments for BPD. However, the lack of a control group and the small sample size in the study limit the generalizability of the finding.

Another gap in the empirical literature on RF in psychotherapy is related to the emotion regulatory role of RF. Many of the positive benefits associated with RF are assumed to be mediated by its role in processing emotional experiences and in tolerating and managing distressing emotions (Fonagy, 2001; Fonagy et al., 1998). However, to our knowledge, the emotion regulatory role of RF in psychotherapy has yet to be examined. Such a role seems highly relevant to BPD, which is often characterized by high levels of negative emotions and with impairments in emotion regulation (Dixon-Gordon et al., 2017). One way to operationalize this emotion regulatory role is to examine if higher quality RF predicts lower emotional arousal. A process in which a therapist bid predicts better patient RF which, in turn, predicts lower patient emotional arousal would suggest that stimulating in-session RF might play a role in in-session emotional processing. Vocally encoded arousal is one promising measure of emotional arousal that may be well suited for this purpose. Vocal pitch as measured by fundamental frequency ( $f_0$ ) may serve as a proxy for emotional arousal because higher vocal pitch is correlated with higher self-reported and psychophysiological arousal (e.g., Baucom et al., 2012).  $f_0$  is an especially promising measure because it is simple to apply, unobtrusive, objective (i.e., does not involve subjective self- or observer-ratings) and is clinically relevant (clinicians are often attuned to fluctuations in vocal pitch and intonations). Indeed,  $f_0$  has been successfully applied to psychotherapy sessions in previous studies showing meaningful associations with indices of emotional processing (Diamond et al., 2010; Imel et al., 2014). Thus, utilizing intensive measurements of therapist bid and patient RF quality in conjunction with vocally encoded arousal may provide a cutting-edge examination of the emotion regulatory role of RF in psychotherapies for BPD.

### The Present Study

The present study is an investigation of the role of in-session RF in a previously completed randomized controlled trial (RCT) of TFP, dialectical behavior therapy (DBT; Linehan, 1993), and supportive psychodynamic therapy (SPT; Appelbaum, 2007) for BPD. RCTs provide an excellent opportunity for research on unique



and common processes because the random assignment of patients to treatment conditions reduces possible confounds related to differences in sample characteristics. Of note, TFP and DBT are two empirically supported (American Psychological Association Division 12, n.d.) and among the most utilized treatment models of BPD (Sansone et al., 2013). TFP is a psychodynamic treatment for BPD that focuses on the patient's disparate representations of themselves and others as they are expressed in the here and now in the relationship with the therapist (Yeomans et al., 2015). In TFP, BPD is conceptualized as a result of unintegrated and undifferentiated self/other representations (i.e., undifferentiated negative representations that are split off from positive ones). Accordingly, the integration of these split-off representations is a presumed principal change mechanism in TFP and it is achieved through increased reflection regarding the disparate aspects of the patient's experience of self and other (Levy, Clarkin, et al., 2006; Yeomans et al., 2015).

DBT is a cognitive-behavioral approach that draws from Buddhist and Dialectical philosophies (Linehan, 1993). It includes four modes of treatment: individual treatment, skills training group, phone coaching for patients, and a consultation team for therapists. DBT balances cognitive and behavioral change strategies with acceptance- and validation-based strategies with the goal of building a life worth living (Linehan, 1993). BPD is conceptualized as a result of biological tendencies toward impulsivity and high emotional sensitivity that are amplified by an invalidating family environment. These risk factors, and the negative consequences associated with them, are presumed to lead to core emotion dysregulation and impairments in behavioral control (Lynch et al., 2006). DBT promotes emotion regulation by reducing maladaptive action tendencies that are associated with intense emotions and by learning how to effectively manage interpersonal conflicts. This, in turn, is achieved through several skills that patients acquire through four skills-training modules (mindfulness, distress tolerance, interpersonal effectiveness, and emotion regulation).

SPT is another approach to the treatment of BPD that has shown preliminary efficacy (Clarkin et al., 2007; reviewed below). It focuses on replacing self-destructive behaviors with verbal expression of conflicts through a healthy collaborative relationship with a therapist, whose reflective capacities serve as a model for the patient. Supportive techniques are utilized to promote identity consolidation and to increase the patient's sense of agency and mastery of impulses and emotions. These techniques may include describing significant aspects of the patient's self, providing emotional support, reassurance, advice, and direct environmental intervention alongside alliance fostering and enhancement of reflective abilities and identification of defensive processes.

The increased RF is presumed to play a central role in the process of change in the TFP model (Levy, Clarkin, et al., 2006) and, to a lesser extent, in the SPT model as well. In contrast, RF is not assumed to play a major role in the DBT model (Lynch et al., 2006). Indeed, a previous study of data drawn from this RCT, showed that RF, as measured during attachment interviews at pre- and posttreatment, increased only during TFP but not during DBT or SPT (Levy, Meehan, et al., 2006).

Thus, the present study builds on these findings to examine in-session RF and its emotion regulatory role. Up to three videotaped sessions per patient, representing early, middle and late phases of treatments, were segmented to talk-turns. Then, talk-turns were

coded for therapist use of bids, patient RF quality, and patient vocally encoded arousal to examine the following hypotheses.

### ***Group Differences and Changes in Study Variables During Treatment (Hypothesis 1)***

Hypothesis 1a: Given the role bids for reflection play in TFP, we hypothesized that the use of bids will be higher in TFP compared to DBT and SPT, especially in the middle phase which represents the working phase of treatment. Hypothesis 1b: similar to and consistent with previous findings (Fischer-Kern et al., 2015; Levy et al., 2021; Levy, Meehan, et al., 2006), we hypothesized that RF quality will increase during TFP but not during DBT and SPT. Research Question 1c: changes in vocally encoded arousal were examined without a specific *a priori* hypothesis.<sup>1</sup>

### ***Within-Session Analyses: the Emotion Regulatory Role of RF (Hypothesis 2)***

Hypothesis 2a: across treatments, therapist bids will predict better patient RF immediately following the bid, which, in turn, will predict lower arousal. Hypothesis 2b: the emotion regulatory effect of RF would increase during treatment, and especially during TFP.

## **Method**

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study. The present study was not preregistered.

### **Participants**

Data were drawn from a previously completed RCT of three treatments for BPD. For a fuller description of the trial, see Clarkin et al. (2007) and Levy, Meehan, et al. (2006). Participants from the New York City Metropolitan area were referred mostly by mental health professionals. Participants were 90 patients who met DSM-IV (American Psychiatric Association, 1994) criteria for primary BPD. Patients with comorbid schizophrenia, schizoaffective disorder, bipolar I disorder, delusional disorder, and/or delirium, dementia, and amnesia and other cognitive disorders were excluded because of the influence of brain pathology and thought disorder on the ability to provide meaningful self-report data and complicated response to treatment. Written informed consent was obtained after all study procedures had been explained, including an agreement to have all assessments and psychotherapy sessions videotaped. The study was approved by the human participant institutional review board. Of the 90 patients, 88 patients had at least one videotaped individual session available for inclusion in the present study. Participants were randomized to one of the three treatment conditions for a no-cost 1-year outpatient treatment. The final sample consisted of 30 TFP, 29 DBT, and 29 SPT patients. The mean age was 31 (range: 19-52, *SD* = 7.9) and the sample was predominantly

<sup>1</sup> Although we did not have an *a priori* directional hypothesis regarding change in vocally encoded arousal, we later became aware of studies showing a decrease in arousal during psychotherapy (e.g., Baucom et al., 2015). In light of these studies, a directional hypothesis might have been more appropriate. We thank an anonymous reviewer for pointing out this literature.



women (93%), Caucasian (68%), unmarried (91%), and educated (58% completed college). The three groups did not differ on any of the demographic variables ( $ps > .05$ ).

## Treatments and Therapists

### *Transference-Focused Psychotherapy (Yeomans et al., 2015)*

Transference-focused psychotherapy is a modified manualized psychodynamic treatment for BPD. Its primary goal is to reduce symptoms and self-destructive behavior through the modification of representations of self and others as they manifest in the treatment. The primary focus is on the predominant affectively charged themes that emerge in the relationship between patients and therapists in sessions.

### *Dialectical Behavior Therapy (Linehan, 1993)*

DBT is a manualized cognitive-behavioral treatment with four modes of treatment: (a) individual therapy, (b) group skills training, (c) phone coaching for patients, and (d) consultation teams for therapists. The primary technique in the individual treatment mode is maintaining a balance of validation and change strategies with the goal of creating a life worth living. The individual treatment focuses on a hierarchy of target behaviors, which the patient tracks daily, often with the use of diary cards. Behavioral analyses of the pattern and chain of thoughts, emotions, and events resulting in suicidal and self-mutilating acts take place routinely to help the patient identify triggers and alternative strategies for coping. In addition, in the group skills training, patients acquire coping skills for reducing maladaptive action tendencies through four modules (mindfulness, distress tolerance, interpersonal effectiveness, and emotion regulation). In the individual treatment, therapist and patient address obstacles in implementing these skills.

### *Supportive Psychodynamic Therapy (Appelbaum, 2007)*

SPT is a psychodynamically oriented treatment for BPD with the primary goal of supporting the patient's adaptive functioning, that is, to bring about changes by developing a healthy collaborative relationship with the therapist and replacing self-destructive enactments with verbal expression of conflicts. Change is assumed to occur by the patient using the therapist's capacity for reflection as a model, as opposed to TFP in which interpretation is used to achieve these goals. SPT utilizes supportive techniques, such as describing significant aspects of the patient's self to encourage greater integration of identity, promoting the patient's sense of agency and self-regulation, and encouraging socially acceptable ways of expressing impulses (such as exercise or creative expression). Emotional support, reassurance, advice, and direct environmental intervention are provided alongside alliance fostering and enhancement of reflection and identification of defensive processes.

Therapists in each of the three treatment conditions were selected based on prior demonstration of competence in their respective treatment and after being judged by cell leaders to be adherent to and competent in their respective manual. All TFP, DBT, and SPT therapists were experienced with patients with

BPD and in their respective modalities. All treatments were supervised weekly by experts to ensure adherence and competence. Throughout the study, all therapists regularly videotaped their sessions and were supervised in a group on a weekly basis. In the subset of 88 patients in the present study, TFP was delivered by 8 therapists that treated an average of 3.75 patients ( $SD = 2.25$ , range: 1–7), DBT was delivered by 4 therapists that treated an average of 7.25 patients ( $SD = 2.99$ , range: 3–10), and SPT was delivered by 7 therapists that treated an average of 4.14 patients ( $SD = 0.69$ , range: 3–5).

## Procedure

Three sessions were identified for each patient, representing three phases of treatment: early (4 months into treatment), middle (8 months) and late (12 months). These specific time points were selected for several reasons. First, we wanted to obtain data that would be most representative of the entire course of treatment. Four months were chosen as the earliest session because there was enough time into the treatment for the sessions to be representative of the working phase of the treatments, and at the same time not too late for disentangling the process from the outcome (i.e., process of late sessions may be a reflection of how successful the outcome of the treatment is). Finally, for future process–outcome analyses, we selected time points that corresponded with the 4-, 8- and 12-month outcome assessments. Overall, 205 recordings of individual therapy (DBT group skills training sessions were not coded) were available for coding: 88 from the early phase, 63 from the middle phase, and 54 from the late phase of treatment. Missing sessions were either due to dropout or sessions not being recorded.

Prior to coding, sessions were segmented into therapist and patient talk-turns. Talk-turns included all the utterances a speaker made from the beginning of the talk-turn to the transition of the turn to the other speaker. Segmentation was done by research assistants using oTranscribe software (<https://otranscribe.com/>) by inserting timestamps at the beginning of each talk-turn and noting whether the speaker was the therapist or the patient. These time stamps were used to make sure that each of the process measures is applied to the exact same talk-turns.

## Measures

### *Therapist Bids for Reflection (Möller et al., 2017)*

Each therapist talk-turn was classified as either asking the patient to consider mental states directly or indirectly ("bid for reflection") or not ("nonbid"). As Möller et al. (2017) mention, an example of a bid would be "why do you think your boyfriend said that?" whereas an example of a nonbid would be "tell me more about what you did around that time." Coding was done by four undergraduate research assistants that were trained to reliability by the first author on a separate sample of patients with BPD (Clarkin et al., 2001). Coders met weekly with the first author to discuss ratings to maintain reliability and avoid rater drift. Coders were blind to study hypotheses, the treatment approach, and timing of the session. A subset of 20 sessions was coded by two coders for interrater reliability calculations. At the session level, interrater reliability of the percent of bids at each session was calculated using a generalizability coefficient for designs in which raters and ratees are neither fully crossed nor nested



(Putka et al., 2008). The intraclass correlation (ICC)-like coefficient was .84. At the statement (within-session) level, absolute agreement was 86% on average with an average Kappa of .55 across sessions. The lower reliability at the within- compared to the between-session level could be due to a restricted range of scores at the within-session level, which has an attenuating effect on interrater reliability estimates (LeBreton et al., 2003; J. LeBreton, personal communication, March 12, 2021).

### **Reflective Functioning Scale (Fonagy et al., 1998)**

An 11-point scale that evaluates the quality of RF. The scale has been mostly applied to the Adult Attachment Interview (AAI; George et al., 1985). We applied the scale to individual patient talk-turns following previous studies (e.g., Möller et al., 2017). The scale ranges from -1 (*negative reflective functioning, overly concrete, totally barren of reflection, or grossly distorting of the mental states of others*) to 9 (*exceptional reflective functioning, unusually complex, elaborate, or original reasoning about mental states*). In addition to coding individual talk-turns, coders also assigned a global score for each session using guidelines that are similar to those in the manual. Coders were four graduate students (i.e., not the same coders that coded bids for RF) and the first author that were trained by the second author and have passed a reliability test by the Anna Freud Center. Every effort was made to keep coders blind to the treatment approach and timing of the session. The graduate student coders, who performed most of the coding (86%) were also blind to study hypotheses. Coders met weekly with the first author to discuss ratings to maintain reliability and avoid rater drift. A subset of 20 sessions was coded by two coders for interrater reliability calculations. At the session level, the ICC-like coefficient (Putka et al., 2008) for global RF scores was .81. In addition, the global scores of the 4-month session correlated positively ( $r = .34, p = .008, n = 60$ ) with reflective functioning scores that were coded from AAI interviews at pre-treatment (Levy, Meehan, et al., 2006). At the statement level, the average ICC across sessions was .62. Similar to the measurement of bids, the lower reliability at the within- compared to the between-session level could be due to a restricted range of scores at the within-session level (LeBreton et al., 2003; J. LeBreton, personal communication, March 12, 2021).

### **Vocally Encoded Arousal**

Vocal pitch as measured by fundamental frequency ( $f_0$ ) served as a measure of emotional arousal. Higher levels of  $f_0$  are indicative of higher pitch and, therefore, higher psychophysiological arousal.  $f_0$  has been successfully applied to psychotherapy sessions in previous studies showing meaningful correlations with indices of emotional processing (Diamond et al., 2010; Imel et al., 2014). First, research assistants cleaned each recording by silencing background noises and overlapping speech. For each patient talk-turn,  $f_0$  was extracted every 0.25 s using Praat speech signal processing software, with a band-pass filter of 75–350 Hz (typical values for human speech; Boersma & Weenink, 2019). The talk-turn  $f_0$  was calculated as the median of all measurements in the talk-turn.

## **Data Analytic Approach**

### **Session Level Analyses: Changes in Study Variables During Treatment (Hypothesis 1)**

Data were analyzed using linear multilevel models implemented in R version 3.4.3, package 'nlme' (Pinheiro et al., 2016). These models account for the multilevel structure of the data (talk-turns repeated within sessions, repeated within patients) while using all available observations, thus providing full intent-to-treat estimates. Analyses that included therapists as an additional level showed that the therapist effect on RF was negligible and nonsignificant (ICC = .00). Therapist effects on bids (ICC = .44) and arousal (ICC = .14) were not negligible but the inclusion of the therapist level did not change the results of these models. Because of that, and because three- and four-level models are difficult to interpret, we only report models that did not include the therapist level. Analyses were adjusted for repeated measures with restricted maximum likelihood estimation method and random intercepts and slopes at the patient level. Diagonal covariance structure was modeled at the patient level and the covariance structure at the session and talk-turn levels (first-order autoregressive, diagonal, or identity) was determined using model fit and deviance tests.

To examine changes in bids for RF, we used the proportion of bids out of the total number of talk-turns in each session as a dependent variable in a two-level model (sessions repeated within patients). Time point (early, middle and late phases of treatment) was included as a Level 1 predictor, treatment group (TFP, DBT, or SPT) was included as a Level 2 predictor, and Group  $\times$  Time point as a cross-level interaction term. Time point was centered around the early phase session so intercepts represent the estimated proportion of bids at that session, while slopes represent the estimated change in the proportion of bids between time points.

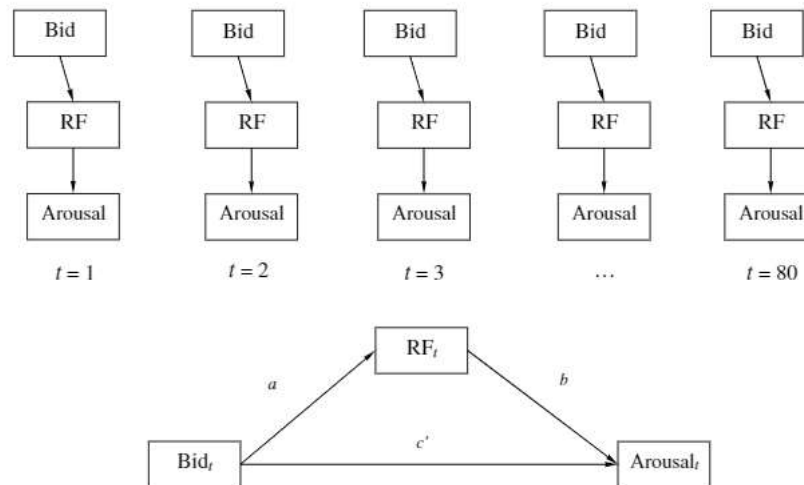
To examine changes in RF and arousal, we fitted three-level models (talk-turns repeated within sessions repeated within patients). Time point (early, middle and late phases of treatment) was included as a Level 2 predictor, treatment arm (Group: TFP, DBT, or SPT) was included as a Level 3 predictor, and Group  $\times$  Time point as a cross-level interaction term. Time point was centered around the early phase session so intercepts represent estimated RF/arousal at that session, while slopes represent estimated change in RF/arousal between time points.

### **Within-Session Analyses: The Emotion Regulatory Role of RF (Hypothesis 2)**

The proposed within-session model was fitted using the dynamic structural equation modeling (DSEM) module in Mplus, version 8.1 (Hamaker et al., 2018; Muthén & Muthén, 2017). A full description of these analyses is included in the [supplemental material, section 2](#). The dataset was structured such that each talk-turn consisted of a therapist intervention (classified as either a bid or a nonbid) and the patient response that *immediately followed the therapist intervention* and was coded for RF and arousal (see [Figure 1](#), top panel, for an illustration). Thus, by design, within each talk-turn, therapist interventions preceded patient responses. Then, a model was fitted that examined all possible associations between the three in-session process variables (therapist bids, patient RF, patient arousal), both concurrently (within a talk-turn) and in a cross-lagged manner.



**Figure 1**  
*Structure of the Within-Session Data (Top Panel) and Hypothesized Mediation Model (Bottom Panel) in the Study*



*Note.* Each talk-turn ( $t$ ) represents a therapist intervention and a subsequent patient response. In the mediation model, therapist bids for reflection are hypothesized to predict better patient reflective functioning (RF) which, in turn, predicts lower patient vocally encoded arousal.

Given the structure of the data, we were mainly interested in the associations within talk-turns. However, modeling the cross-lagged associations allows to take into account the state of affairs in the previous talk-turn in a dynamic manner.

To test the hypothesis that bids will be associated with better RF which, in turn, will be associated with lower arousal we examined a mediation model (Figure 1, bottom panel). We estimated the path from bids to RF immediately following the bid (called the  $a$  path in the mediation literature) and the path from RF to arousal at the same talk-turn (called the  $b$  path). The  $a \times b$  term was then estimated as a measure of the indirect effect. Our hypothesis was that better RF should have a down-regulatory effect on arousal as it occurs *at the moment*. We, therefore, focused on the path from RF to arousal at the same talk-turn and not in a lag. In addition, by design, RF at a given talk-turn and arousal at the subsequent talk-turn are separated by a therapist bid and, therefore, examining the cross-lagged association in the mediation model would not allow to disentangle the effect of previous RF and the therapist bid on arousal.

The model was first examined on all three sessions combined. To examine changes over time in the indirect path, we analyzed person-specific estimates of the indirect path using multilevel modeling using the approach that was taken to examine changes in study-level variables.

Finally, to strengthen the validity of our hypothesized model, we also tested a competing model. We examined whether lower arousal predicts better RF rather than the other way around (e.g., whether being less emotionally aroused allows patients to better utilize RF). To examine this, we tested the indirect effect of bids on better RF through lower arousal.

The data and study materials are not available as they include personally identifiable information. The data analysis code is available upon request from the first author.

## Results

### Descriptive Statistics

The final dataset in the present study included 88 patients and 205 sessions (59 sessions were not available for coding due to patient dropout or missing recordings). In total, those sessions included 17,973 therapist-patient talk-turns (each consisting of a therapist statement and a subsequent patient statement) and 2,232,333 observations of vocal data (that were aggregate using the median of each statement). TFP had less talk-turns per session ( $M = 65.84$ ,  $SD = 32.40$ ) compared to DBT ( $M = 102.21$ ,  $SD = 39.22$ ;  $t_{108} = 5.67$ ,  $p < .001$ ) and SPT ( $M = 98.31$ ,  $SD = 32.95$ ;  $t_{146} = 6.05$ ,  $p < .001$ ), which did not differ from one another ( $t_{109} = 0.60$ ,  $p = .55$ ).

### Session Level Analyses: Changes in Study Variables During Treatment

Session-level means, and standard deviations for therapist bids (ratio of bids out of all talk-turns in the session), patient RF following a bid, patient RF following a nonbid, and patient vocally encoded arousal ( $f_0$ ) are reported in Table 1, separately for group and time point.

### Change in Therapist Bids for RF

A multilevel model examined the effect of time (early, mid, and late therapy) and treatment group on the ratio of therapist bids for RF per session. The results showed a significant main effect of group,  $F_{(2, 85)} = 37.16$ ,  $p < .001$ , such that TFP therapists had a higher ratio of bids ( $M = .35$ ,  $SD = .15$ ) than DBT therapists,  $M = .17$ ,  $SD = .07$ ;  $F_{(1, 109)} = 52.16$ ,  $p < .001$ , and SPT therapists,  $M = .17$ ,  $SD = .08$ ;  $F_{(1, 109)} = 56.32$ ,  $p < .001$ . DBT and SPT therapists did not differ

**Table 1**

*Session-Level Means and Standard Deviations of Bids for RF, RF Following a Bid, RF Following a Nonbid and Vocally Encoded Arousal by Treatment Group and Time point*

Measure/time point	TFP	DBT	SPT
Bids for RF (ratio)			
4-month session	0.36 (0.20)	0.19 (0.10)	0.18 (0.11)
8-month session	0.40 (0.17)	0.14 (0.09)	0.19 (0.10)
12-month session	0.33 (0.16)	0.15 (0.07)	0.13 (0.08)
RF following a nonbid			
4-month session	1.91 (0.68)	1.78 (0.47)	1.92 (0.47)
8-month session	2.17 (0.77)	2.00 (0.43)	1.97 (0.45)
12-month session	2.16 (0.82)	1.86 (0.52)	1.89 (0.47)
RF following a bid			
4-month session	2.25 (0.96)	2.29 (0.85)	2.31 (0.63)
8-month session	2.39 (1.01)	2.67 (0.60)	2.41 (0.75)
12-month session	2.50 (0.87)	2.44 (0.78)	2.39 (0.67)
Vocally encoded arousal			
4-month session	167.91 (25.50)	176.37 (31.77)	181.43 (27.81)
8-month session	159.54 (27.50)	174.78 (23.58)	184.48 (25.59)
12-month session	168.20 (30.20)	174.16 (28.17)	188.69 (28.36)

*Note.* TFP = transference-focused psychotherapy; DBT = dialectical behavior therapy; SPT = supportive psychodynamic therapy; RF = reflective functioning.

from one another,  $F_{(1, 109)} = 0.10, p = .76$ . Thus, TFP therapists made bids for RF in 35% of their interventions while DBT and SPT therapists showed half the ratio of bids (17%) compared to TFP. In addition, no significant effects emerged for time,  $F_{(2, 111)} = 2.04, p = .13$ , and Time  $\times$  Group,  $F_{(2, 111)} = 0.63, p = .64$ , which means that the ratio of bids did not change during treatment in any of the groups.

### Change in Patient In-Session RF

As Table 1 shows, on average, patient RF tended to be rather low in all treatments and time points (average scores were around 2 while scores of 3 are considered questionable or low RF according to the RF scale). A multilevel model examined the effect of time and treatment group on RF. No significant effects emerged for group,  $F_{(2, 85)} = 1.27, p = .29$ , time,  $F_{(2, 111)} = 1.30, p = .28$ , and Time  $\times$  Group,  $F_{(2, 111)} = 0.42, p = .79$ . Thus, no changes in RF were observed during any of the treatments.

### Change in Patient Vocally Encoded Arousal

A multilevel model examined the effect of time and treatment groups on patient arousal. The results showed a significant main effect of group,  $F_{(2, 85)} = 3.66, p = .03$ , such that patients in TFP had lower arousal ( $M = 165.77, SD = 22.67$ ) compared to patients in SPT,  $M = 183.32, SD = 25.10$ ;  $F_{(1, 109)} = 7.10, p = .009$ . Patients in DBT ( $M = 177.20, SD = 32.17$ ) did not differ from patients in TFP,  $F_{(1, 109)} = 2.84, p = .09$ , or from patients in SPT,  $F_{(1, 109)} = 0.81, p = .37$ . No significant effects emerged for time,  $F_{(2, 111)} = 2.20, p = .12$ , and Time  $\times$  Group,  $F_{(2, 111)} = 1.89, p = .12$ , which means that arousal did not change during treatment in any of the groups.

### Within-Session Analyses: The Emotion Regulatory Role of RF

Figure 1 shows the structure of the within-session data (top panel) and the hypothesized mediation model (bottom panel). The SEM fitted to test the study hypotheses across all three sessions is displayed in Figure 2. It is important to note that the model was fitted and applies to all available talk-turns in each session, but only two adjacent talk-turns are presented for ease of interpretation. All model paths express the effect of that path while taking into account all other paths in the model. Unstandardized estimates (denoted  $b$ ) represent the expected change in the outcome variable that is associated with a one-point increase in the predictor. Standardized estimates (denoted  $\beta$ ) represent the expected change in standard deviations in the outcome variable that is associated with an increase of one standard deviation in the predictor.

### The Emotion Regulatory Role of RF

Consistent with our hypotheses, across sessions, we found that bids predicted better RF immediately following the bid,  $b = 0.17$  [ $0.12, 0.22$ ],  $p < .001, \beta = .15$  [ $.12, .17$ ]. This means that patient RF tended to be 0.17 points higher on talk-turns that were preceded by a therapist bid than on talk-turns that were preceded by a nonbid. In addition, better RF predicted lower vocally encoded arousal at the same talk-turn,  $b = -0.94$  [ $-1.29, -0.59$ ],  $p < .001, \beta = -.05$  [ $-.07, -.04$ ]. Importantly, the indirect effect of therapist bids on patient vocally encoded arousal through better patient RF was significant,  $b = -0.15$  [ $-0.24, -0.08$ ],  $p < .001$ . Therefore, an emotion regulatory effect RF was observed such that use of bids predicted better patient RF which, in turn, predicted lower emotional arousal.

### Additional Paths in the Model

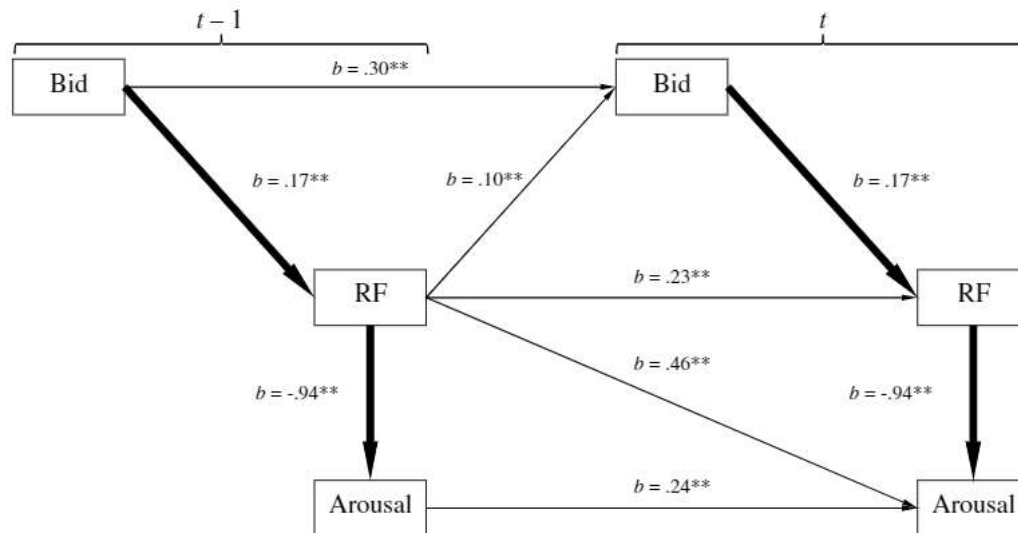
In addition to findings above, additional paths emerged as significant in the model. All autoregressive paths in the model were significant suggesting that a moderate degree of stability in the likelihood of the occurrence of bids,  $b = 0.30$  [ $0.24, 0.35$ ],  $p < .001, \beta = .29$  [ $.26, .33$ ], in patient RF,  $b = 0.23$  [ $0.19, 0.27$ ],  $p < .001, \beta = .23$  [ $.21, .24$ ], and in patient vocally encoded arousal,  $b = 0.24$  [ $0.19, 0.28$ ],  $p < .001, \beta = .24$  [ $.22, .26$ ]. In addition, better RF predicted greater likelihood of subsequent bids,  $b = 0.10$  [ $0.06, 0.14$ ],  $p < .001, \beta = .11$  [ $.08, .13$ ], and higher subsequent arousal,  $b = 0.46$  [ $0.12, 0.80$ ],  $p = .01, \beta = .02$  [ $.01, .04$ ]. Bids did not directly predict lower arousal immediately following the bid (after taking into account the indirect effect through better RF;  $b = 0.39$  [ $-0.20, 1.04$ ],  $p = .19, \beta = .02$  [ $-.01, .05$ ]). Furthermore, bids did not predict arousal,  $b = -0.44$  [ $-1.04, 0.13$ ],  $p = .13, \beta = .02$  [ $-.05, .01$ ], and RF,  $b = 0.04$  [ $-0.01, 0.08$ ],  $p = .16, \beta = .03$  [ $.00, .05$ ], in the talk-turn that followed the next bid (e.g., two bids apart). Finally, lower arousal did not predict greater likelihood of subsequent bids,  $b = 0.00$  [ $-0.02, 0.02$ ],  $p = .88, \beta = -.03$  [ $-.05, .00$ ], or better subsequent RF,  $b = 0.00$  [ $-0.02, 0.02$ ],  $p = .97, \beta = .00$  [ $-.02, .02$ ].

In sum, bids predicted lower arousal through better RF. In addition, RF predicted a greater likelihood of additional bids in the subsequent talk-turn and also higher arousal. The increased subsequent RF, in turn, predicted lower subsequent arousal. Bids,



**Figure 2**

*A Dynamic Structural Equation Model Examining the Emotion Regulatory Effect of Bids for Reflection on Lower Vocally Encoded Arousal Through Better Reflective Functioning (RF)*



*Note.* Only significant paths are presented. Indirect path is presented by thick arrows. The model was applied to all available talk-turns but for ease of interpretation, only two adjacent talk-turns are presented.

$^{**} p < .001$ .

RF, and arousal showed a moderate degree of stability between talk-turns.

### **Group Differences and Changes During Treatment in the Emotion Regulatory Role of RF**

To examine changes during treatment, a multilevel model was fitted to session-level estimates of the indirect effect of bids on arousal through RF (i.e., an estimate of the indirect path was saved separately for each patient and each session from the SEM). We found that the indirect effect at the 4-month session was significant ( $b = -0.14$ ,  $t_{114} = -13.88$ ,  $p < .001$ ) and it became stronger over time ( $b = -0.05$ ,  $t_{114} = -4.46$ ,  $p < .001$ ). The effects of group,  $F_{(2, 85)} = 0.07$ ,  $p = .93$ , and Time  $\times$  Group,  $F_{(2, 114)} = 0.47$ ,  $p = .63$ , were not significant.

To better understand this finding, we examined changes during treatment in each of the paths that constitute the indirect effect, namely the effect of bids on RF immediately following the bid and the effect of RF on arousal. In a model examining the effect of bids on RF immediately following the bid, none of the effects were significant, time:  $F_{(2, 114)} = 0.01$ ,  $p = .93$ ; group:  $F_{(2, 85)} = 0.02$ ,  $p = .98$ ; Time  $\times$  Group:  $F_{(2, 114)} = 0.38$ ,  $p = .68$ . However, the effect of RF on decreasing arousal became stronger over time ( $b = -0.26$ ,  $t_{114} = -14.00$ ,  $p < .001$ ) while the group,  $F_{(2, 85)} = 1.40$ ,  $p = .25$ , and Time  $\times$  Group,  $F_{(2, 114)} = 1.71$ ,  $p = .19$ , effects were not significant. Thus, these findings suggest that the regulatory effect of RF on arousal became stronger during treatment and that this increase could be attributed to an increase during treatment in the effect of RF on decreased arousal. In addition, the regulatory effect did not differ by group.

### **Testing a Competing Model: Lower Arousal as a Facilitator of Better RF**

Given the naturalistic nature of the within-session model, it was important to test plausible competing models to the hypothesized model. Therefore, we tested a model with an indirect effect of bids on better RF through lower arousal. The model examined the alternative explanation that lower arousal predicts better RF and not the other way around. This might be the case if being less emotionally aroused allows patients to better utilize RF. We found that bids did not predict lower subsequent arousal,  $b = 0.23$  [ $-0.32$ ,  $0.81$ ],  $p = .37$ ,  $\beta = .07$  [ $-.05$ ,  $.23$ ], and arousal did not predict better RF,  $b = 0.00$  [ $-0.03$ ,  $0.02$ ],  $p = .82$ ,  $\beta = .00$  [ $-.01$ ,  $.01$ ]. Similarly, the indirect effect of bids on RF through arousal was not significant,  $b = 0.00$  [ $-0.02$ ,  $0.01$ ],  $p = .86$ . These findings provide further support to the proposed model by ruling out the possibility that lower arousal contributes to better RF and not vice versa.

### **Accounting for Therapist Vocally Encoded Arousal**

A possible alternative explanation for the indirect effect of bids on arousal through RF is that therapist arousal serves as a third variable that is correlated with increased likelihood of bids and also predicts lower subsequent patient arousal. For example, making a bid might reduce the therapist's arousal. In turn, therapist arousal may predict lower patient arousal, for example, by interpersonal regulation/dampening (Imel et al., 2014). To rule out this explanation, we tested a model that included therapist arousal. We examined whether bids predicted therapist arousal in the same talk-turn and whether patient arousal was predicted from therapist arousal. The results showed that bids indeed predicted lower therapist arousal in the



same talk-turn,  $b = -0.93$   $[-1.47, -.37]$ ,  $p < .001$ , and that therapist arousal predicted lower patient arousal,  $b = 0.11$   $[0.07, 0.16]$ ,  $p < .001$ . However, the indirect effect of bids on patient arousal through RF remained significant,  $b = -0.15$   $[-0.24, -0.08]$ ,  $p < .001$ . Thus, although an indirect path between bids and patient arousal through therapist arousal may exist, it seems that the emotion regulatory effect of bids through RF is independent of that path.

## Discussion

Impairments in RF, the capacity to consider mental states, are considered central in the psychopathology of BPD, and their reduction is hypothesized to play a role in the treatment of BPD, and especially in psychodynamic psychotherapies (Levy, Clarkin, et al., 2006). The present study examined the emotion regulatory role of in-session RF in psychotherapies for BPD. Findings from this intensive, microanalytic research design showed that the therapist interventions that ask participants to consider mental state are associated with better subsequent patient RF which, in turn, predicts reduced emotional arousal. Below we discuss these findings in detail.

### The Emotion Regulatory Role of In-Session RF

Consistent with our hypothesis and with a previous study that utilized a similar design (Möller et al., 2017), we found that bids predicted better RF in the talk-turn immediately following the bid. This finding suggests that asking patients to consider mental states results in a better quality of RF. Furthermore, our findings showed that better RF predicted lower vocally encoded arousal—suggesting that RF may play a role in downregulating emotional arousal in session. Although the effect size tended to be small, this finding is consistent with the notion of RF as enabling meaning-making that promotes processing of interpersonal and emotional experiences thereby increasing tolerance and regulation of distressing mental states (Fonagy, 2001). To our knowledge, this is the first study to show that RF as measured by an observer predicts a psychophysiological measure of arousal in the context of psychotherapy.

It is noteworthy that our findings held even after examining (and ruling out) alternative explanations such as patient arousal predicting RF and not the other way around or the possibility that therapist arousal, and not RF, predicts lower patient arousal. That said, another possible alternative explanation for our findings is that patients experience bids as “demands” for RF, and having complied with the therapists’ demands, patients feel a relief which is expressed in lower arousal.

It is important to note that the direct effect of bids on arousal (that is, not including the indirect effect through better RF) was not significant. This finding may suggest that bids for RF do not have a universal down-regulatory effect on arousal. Rather, bids are associated with decreased arousal only to the degree that they are also associated with increased RF. It remains for future studies to try to identify what makes for a successful bid. For example, optimal bids might be those that require a moderate level of incremental reflection from the patient while trivial or overly demanding bids might not be successful in promoting better RF and lower arousal.

Taken together, we see the indirect effect of bids on arousal through RF as capturing the degree to which patients can take in

what the therapist is saying and use it to reflect on their experience to downregulate their distressing emotions. Such a process plays a central role in many psychodynamic therapies and is an important part of the TFP model (Levy, Clarkin, et al., 2006). These findings demonstrate the contribution of microanalytic designs coupled with appropriate novel data analytic approaches to understanding the moment-by-moment unfolding of in-session processes. Future studies could further examine the causal effect of bids on RF and arousal in more controlled settings by utilizing experimental designs such as lab experiments and microinterventions. Future studies could also examine a more nuanced classification of bids as well as take into account the competence in which the bid is delivered. For example, from a TFP perspective, studies could examine differences between interventions such as clarifications, confrontations, and interpretations in their effects on subsequent RF.

### Group Differences and Changes Over Time

Consistent with hypotheses and with the TFP model, bids were twice as common in TFP compared to DBT and SPT. Use of bids did not increase during treatment in TFP, in contrast to hypotheses, which means that bids were common even early in treatment, at least during the active phase of it (i.e., after goal- and frame setting).

Interestingly, the emotion regulatory effect of bids was similar regardless of treatment despite the fact that RF and bids for RF are not an integral part of the DBT model (Lynch et al., 2006) and play a lesser role in the SPT model compared to TFP (Appelbaum, 2007). Thus, bids were more likely to occur in TFP, but when they occurred, they had a similar effect on RF and arousal regardless of treatment.

In contrast to hypotheses, in-session RF rated on therapy discourse did not increase during treatment, including in TFP. This finding is inconsistent with a previous study of the same dataset that showed increases in RF during TFP only, from pre- to posttreatment, as measured by an attachment interview (the AAI). There are several possible explanations for this discrepancy in findings. TFP patients may have improved in their capacity to reflect on their past experiences and/or with regard to their parents (aspects that are captured by the AAI) whereas the capacity to reflect at the moment and/or with regard to the therapist/significant others during psychotherapy sessions is slower to develop. Related to that, current upsetting experiences that are discussed in session may be more emotionally evoking, and, therefore, harder to reflect on, compared to talking about past experiences. Examining RF at follow-up periods after treatment as well as using methods for assessing RF in daily life between sessions (e.g., ecological momentary assessments) could help clarify the pathways for improvements in RF during TFP.

### Study Limitations and Future Research

Psychotherapy, by definition, occurs in an interpersonal context. Therefore, the findings in the present study should be interpreted as reflecting dyadic processes, rather than pure patient-related processes. For example, it is impossible to determine if RF, as coded in the study, solely reflects the patient’s capacity for reflection rather than interaction between the patient’s capacity to reflect and the therapist’s capacity to facilitate reflection (mediated by the therapist’s own RF capacity). As mentioned above, coding the skill in



which bids were delivered, as well as therapists' capacity to reflect on their patients' experiences may help disentangle these contributions (Diamond et al., 2003).

In the present study, we focused on emotional arousal, broadly defined, regardless of the content and valence of the emotion. Although it is reasonable to assume that on most talk-turns emotional arousal was related to unpleasant emotions, we cannot rule out that some talk-turns included emotional arousal related to pleasant emotions. The process described in the study may be different for pleasant emotions and, therefore, future studies could complement our findings by classifying talk-turns based on their affective content.

Related to that, it is important to note that we do not assume that decreasing arousal is a uniform goal in psychotherapy that applies in all situations. Indeed, some studies show that a moderate level of arousal tends to predict the best outcomes (Carryer & Greenberg, 2010). Instead, our goal was to examine RF-related reductions in arousal that may be a proxy for emotional processing. Given that individuals with BPD are more likely to be dysregulated, rather than constricted, in their emotional experience (Dixon-Gordon et al., 2017), a decrease, rather than an increase, in arousal seemed to be more relevant. Nevertheless, future studies could try to identify for which patients a reduction in arousal is more likely to be conducive to emotional processing and which patients need to experience their emotions more fully to process them. Attachment styles, and their related emotion regulation patterns, may be one factor distinguishing between these two groups of patients (Shaver & Mikulincer, 2007).

The large number of observations per patient (80 talk-turns per session on average and up to three sessions per patient) made our analyses well powered for detecting within-patient processes. However, with a total of 88 patients in three treatments, the study may not have been adequately powered to detect group differences. Therefore, nonsignificant group differences in the present study should be interpreted with caution and replicated in larger samples.

Finally, the present study focused exclusively on in-session processes. As a next step, it would be important to examine how in-session RF and its emotion regulatory effect are related to treatment outcome. Psychotherapy outcome is most often conceived as occurring in longer time intervals than in-session processes, typically at the between-session level (i.e., session by session changes in symptoms). Thus, bridging between these micro- (within-session processes) and mesophenomena (between-session outcome) is an important challenge to address in cutting edge process-outcome research (Elliott, 2010; Penix et al., 2020).

### Implications and Concluding Comments

Psychotherapy is a dynamic process that unfolds moment-by-moment. The present study demonstrates the utility of intensive measurements and of microanalytic research designs in capturing these dynamic processes. The findings of the study highlight the importance of therapists' focus on increasing patient RF, which, in turn, predicts lower emotional arousal. These findings represent a step towards elucidating unique and common mechanisms of change in psychotherapies for BPD, a severe disorder which is of major public health concern. Such findings, coupled with future studies focusing on in-session processes-outcome relations, have

the potential of informing the practice and delivery of psychotherapies for BPD to maximize the emotion regulatory role of RF.

### References

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.).
- American Psychological Association. Division 12. (n.d.). *Psychological treatments*. <https://www.div12.org/treatments/>
- Appelbaum, A. H. (2007). Supportive psychotherapy. In J. M. Oldham, A. E. Skodol, & D. S. Bender (Eds.), *The American Psychiatric Publishing Textbook of Personality Disorders* (pp. 311–326). American Psychiatric Publishing.
- Bateman, A., Campbell, C., Luyten, P., & Fonagy, P. (2018). A mentalization-based approach to common factors in the treatment of borderline personality disorder. *Current Opinion in Psychology*, 21, 44–49. <https://doi.org/10.1016/j.copsyc.2017.09.005>
- Baucom, B. R., Saxbe, D. E., Ramos, M. C., Spies, L. A., Iturralde, E., Duman, S., & Margolin, G. (2012). Correlates and characteristics of adolescents' encoded emotional arousal during family conflict. *Emotion*, 12(6), 1281–1291. <https://doi.org/10.1037/a0028872>
- Baucom, B. R., Sheng, E., Christensen, A., Georgiou, P. G., Narayanan, S. S., & Atkins, D. C. (2015). Behaviorally-based couple therapies reduce emotional arousal during couple conflict. *Behaviour Research and Therapy*, 72, 49–55. <https://doi.org/10.1016/j.brat.2015.06.015>
- Boersma, P., & Weenink, D. (2019). *Praat: doing phonetics by computer [Computer program]*. Version 6.0.55. <http://www.praat.org/>
- Carryer, J. R., & Greenberg, L. S. (2010). Optimal levels of emotional arousal in experiential therapy of depression. *Journal of Consulting and Clinical Psychology*, 78(2), 190–199. <https://doi.org/10.1037/a0018401>
- Clarkin, J. F., Foelsch, P. A., Levy, K. N., Hull, J. W., Delaney, J. C., & Kernberg, O. F. (2001). The development of a psychodynamic treatment for patients with borderline personality disorder: A preliminary study of behavioral change. *Journal of Personality Disorders*, 15(6), 487–495. <https://doi.org/10.1521/pedi.15.6.487.19190>
- Clarkin, J. F., & Levy, K. N. (2006). Psychotherapy for patients with borderline personality disorder: Focusing on the mechanisms of change. *Journal of Clinical Psychology*, 62(4), 405–410. <https://doi.org/10.1002/jclp.20238>
- Clarkin, J. F., Levy, K. N., Lenzenweger, M. F., & Kernberg, O. F. (2007). Evaluating three treatments for borderline personality disorder: A multi-wave study. *The American Journal of Psychiatry*, 164(6), 922–928. <https://doi.org/10.1176/ajp.2007.164.6.922>
- Cristea, I. A., Gentili, C., Cotet, C. D., Palomba, D., Barbu, C., & Cuijpers, P. (2017). Efficacy of psychotherapies for borderline personality disorder: A systematic review and meta-analysis. *JAMA Psychiatry*, 74(4), 319–328. <https://doi.org/10.1001/jamapsychiatry.2016.4287>
- Diamond, G., Rochman, D., & Amir, O. (2010). Arousing primary vulnerable emotions in the context of unresolved anger. *Journal of Counseling Psychology*, 57(4), 402–410. <https://doi.org/10.1037/a0021115>
- Diamond, D., Stovall-McClough, C., Clarkin, J. F., & Levy, K. N. (2003). Patient-therapist attachment in the treatment of borderline personality disorder. *Bulletin of the Menninger Clinic*, 67(3), 227–259. <https://doi.org/10.1521/bumc.67.3.227.23433>
- Dixon-Gordon, K. L., Peters, J. R., Fertuck, E. A., & Yen, S. (2017). Emotional processes in borderline personality disorder: An update for clinical practice. *Journal of Psychotherapy Integration*, 27(4), 425–438. <https://doi.org/10.1037/int0000044>
- Elliott, R. (2010). Psychotherapy change process research: Realizing the promise. *Psychotherapy Research*, 20(2), 123–135. <https://doi.org/10.1080/10503300903470743>
- Fischer-Kern, M., Doering, S., Taubner, S., Hörz, S., Zimmermann, J., Rentrop, M., Schuster, P., Buchheim, P., & Buchheim, A. (2015). Transference-focused psychotherapy for borderline personality disorder:



- Change in reflective function. *The British Journal of Psychiatry*, 207(2), 173–174. <https://doi.org/10.1192/bjp.bp.113.143842>
- Fonagy, P. (2001). The human genome and the representational world: The role of early mother–infant interaction in creating an interpersonal interpretive mechanism. *Bulletin of the Menninger Clinic*, 65(3), 427–448. <https://doi.org/10.1521/bumc.65.3.427.19844>
- Fonagy, P., & Bateman, A. W. (2006). Mechanisms of change in mentalization-based treatment of BPD. *Journal of Clinical Psychology*, 62(4), 411–430. <https://doi.org/10.1002/jclp.20241>
- Fonagy, P., & Luyten, P. (2009). A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder. *Development and Psychopathology*, 21(4), 1355–1381. <https://doi.org/10.1017/S0954579409990198>
- Fonagy, P., Target, M., Steele, H., & Steele, M. (1998). *Reflective-functioning manual, version 5.0* [Unpublished manuscript]. University College.
- George, C., Kaplan, N., & Main, M. (1985). *The Berkeley Adult Attachment Interview* [Unpublished manuscript]. University of California.
- Hamaker, E. L., Asparouhov, T., Brose, A., Schmiedek, F., & Muthén, B. (2018). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study. *Multivariate Behavioral Research*, 53(6), 820–841. <https://doi.org/10.1080/00273171.2018.1446819>
- Imel, Z. E., Barco, J. S., Brown, H. J., Baucom, B. R., Baer, J. S., Kircher, J. C., & Atkins, D. C. (2014). The association of therapist empathy and synchrony in vocally encoded arousal. *Journal of Counseling Psychology*, 61(1), 146–153. <https://doi.org/10.1037/a0034943>
- Kivity, Y., Levy, K. N., Rosenstein, L. K., Johnson, B. N., & LeBreton, J. M. (2021). *Mentalizing in and out of awareness: A conceptual framework and a meta-analytic review of implicit and explicit mentalizing* [Manuscript in preparation].
- Kivity, Y., Levy, K. N., Wasserman, R. H., Beeney, J. E., Meehan, K. B., & Clarkin, J. F. (2019). Conformity to prototypical therapeutic principles and its relation with change in reflective functioning in three treatments for borderline personality disorder. *Journal of Consulting and Clinical Psychology*, 87(11), 975–988. <https://doi.org/10.1037/ccp0000445>
- LeBreton, J. M., Burgess, J. R., Kaiser, R. B., Atchley, E. K., & James, L. R. (2003). The restriction of variance hypothesis and interrater reliability and agreement: Are ratings from multiple sources really dissimilar? *Organizational Research Methods*, 6(1), 80–128. <https://doi.org/10.1177/1094428102239427>
- Levy, K. N., Clarkin, J. F., Yeomans, F. E., Scott, L. N., Wasserman, R. H., & Kernberg, O. F. (2006). The mechanisms of change in the treatment of borderline personality disorder with transference focused psychotherapy. *Journal of Clinical Psychology*, 62(4), 481–501. <https://doi.org/10.1002/jclp.20239>
- Levy, K. N., Diamond, D., Clarkin, J. F., & Kernberg, O. F. (2021). *Changes in attachment, reflective function, and object representation in transference focused psychotherapy for borderline personality disorder* [Manuscript in preparation].
- Levy, K. N., Meehan, K. B., Kelly, K. M., Reynoso, J. S., Weber, M., Clarkin, J. F., & Kernberg, O. F. (2006). Change in attachment patterns and reflective function in a randomized control trial of transference-focused psychotherapy for borderline personality disorder. *Journal of Consulting and Clinical Psychology*, 74(6), 1027–1040. <https://doi.org/10.1037/0022-006X.74.6.1027>
- Linehan, M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. Guilford Press.
- Lynch, T. R., Chapman, A. L., Rosenthal, M. Z., Kuo, J. R., & Linehan, M. M. (2006). Mechanisms of change in dialectical behavior therapy: Theoretical and empirical observations. *Journal of Clinical Psychology*, 62(4), 459–480. <https://doi.org/10.1002/jclp.20243>
- Möller, C., Karlgren, L., Sandell, A., Falkenström, F., & Philips, B. (2017). Mentalization-based therapy adherence and competence stimulates in-session mentalization in psychotherapy for borderline personality disorder with co-morbid substance dependence. *Psychotherapy Research*, 27(6), 749–765. <https://doi.org/10.1080/10503307.2016.1158433>
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide* (8th ed.).
- Penix, E. A., Swift, J. K., & Trusty, W. T. (2020). Integrating clients' moment-to-moment ratings into psychotherapy research: A novel approach. *Counseling & Psychotherapy Research*, 20(3), 456–462. <https://doi.org/10.1002/capr.12296>
- Pinheiro, J., & Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2016). *nlme: Linear and Nonlinear Mixed Effects Models, R package version 3.1-128* [Computer software]. <http://CRAN.R-project.org/package=nlme>
- Putka, D. J., Le, H., McCloy, R. A., & Diaz, T. (2008). Ill-structured measurement designs in organizational research: Implications for estimating interrater reliability. *Journal of Applied Psychology*, 93(5), 959–981. <https://doi.org/10.1037/0021-9010.93.5.959>
- Sansone, R. A., Kay, J., & Anderson, J. L. (2013). Resident didactic education in borderline personality disorder: Is it sufficient? *Academic Psychiatry*, 37(4), 287–288. <https://doi.org/10.1176/appi.ap.12110194>
- Shaver, P. R., & Mikulincer, M. (2007). Adult attachment strategies and the regulation of emotion. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 446–465). The Guilford Press.
- Skodol, A. E., Gunderson, J. G., Pfohl, B., Widiger, T. A., Livesley, W. J., & Siever, L. J. (2002). The borderline diagnosis I: Psychopathology, comorbidity, and personality structure. *Biological Psychiatry*, 51(12), 936–950. [https://doi.org/10.1016/S0006-3223\(02\)01324-0](https://doi.org/10.1016/S0006-3223(02)01324-0)
- Yeomans, F. E., Clarkin, J. F., & Kernberg, O. F. (2015). *Transference-focused psychotherapy for borderline personality disorder: A clinical guide*. American Psychiatric Publishing.

Received December 23, 2020

Revision received July 30, 2021

Accepted August 2, 2021 ■