American Psychological Association

Journal of Consulting and Clinical Psychology

ISSN: 0022-006X

2021, Vol. 89, No. 3, 227–239
https://doi.org/10.1037/ccp0000623

Intrapersonal and Interpersonal Vocal Affect Dynamics During Psychotherapy

Adar Paz[1], Eshkol Rafaeli[1], Eran Bar-Kalifa[2], Eva Gilboa-Schechtman[1], Sharon Gannot[3], Bracha Laufer-Goldshtein[3], Shrikanth Narayanan[4], Joseph Keshet[5], and Dana Atzil-Slonim[1]

[1] Department of Psychology, Bar-Ilan University
[2] Department of Psychology, Ben-Gurion University of the Negev
[3] Faculty of Engineering, Bar-Ilan University
[4] Signal Analysis and Interpretation Laboratory (SAIL), University of Southern California
[5] Department of Computer Science, Bar-Ilan University

Objective: The present study implements an automatic method of assessing arousal in vocal data as well as dynamic system models to explore intrapersonal and interpersonal affect dynamics within psychotherapy and to determine whether these dynamics are associated with treatment outcomes. ***Method:*** The data of 21,133 mean vocal arousal observations were extracted from 279 therapy sessions in a sample of 30 clients treated by 24 therapists. Before and after each session, clients self-reported their well-being level, using the Outcome Rating Scale. ***Results:*** Both clients' and therapists' vocal arousal showed intrapersonal dampening. Specifically, although both therapists and clients departed from their baseline, their vocal arousal levels were "pulled" back to these baselines. In addition, both clients and therapists exhibited interpersonal dampening. Specifically, both the clients' and the therapists' levels of arousal were "pulled" toward the other party's arousal level, and clients were "pulled" by their therapists' vocal arousal toward their own baseline. These dynamics exhibited a linear change over the course of treatment: whereas interpersonal dampening decreased over time, there was an increase in intrapersonal dampening over time. In addition, higher levels of interpersonal dampening were associated with better session outcomes. ***Conclusions:*** These findings demonstrate the advantages of using automatic vocal measures to capture nuanced intrapersonal and interpersonal affect dynamics in psychotherapy and demonstrate how these dynamics are associated with treatment gains.

Public Health Significance Statement
The current findings highlight the potential of computerized vocal analyses to capture moment-by-moment processes within psychotherapy sessions. They suggest that clients and therapists exhibit both intrapersonal (within person) as well as interpersonal (between person) affect dynamics in their in-session emotional arousal levels. Specifically, both clients and therapists not only tended to return to their own affective arousal baseline but also tended to be "pulled" by their partner toward *their* baseline arousal level. The findings advance the idea that therapists who are synchronized with their clients, but at the same time downregulate their own and their clients' affect, may be more successful in helping their clients develop better affective regulation capabilities.

Keywords: voice analysis, intrapersonal and interpersonal affect dynamics, coregulation, self-regulation, vocal arousal

Disrupted affect regulation processes are posited to be at the epicenter of many mental disorders (Joormann & Stanton, 2016; Sheppes et al., 2015), and the modification of these processes is at the core of many therapeutic interventions. Understanding affect and affective arousal within psychotherapy necessitates an analysis of how these fluctuate and change over time within the client (i.e., intrapersonal affect dynamics) as well as between the client and the therapist (i.e., interpersonal affect dynamics) and the extent to which these dynamics are associated with treatment outcomes (Greenberg, 2012; Fosha, 2001).

To date, temporal dynamics in affective arousal during therapy have typically been modeled on clients' subjective reports completed retrospectively for entire sessions (e.g., Atzil-Slonim et al., 2018; Fisher et al., 2016). Moreover, most of these studies have focused solely on the client's experience, eschewing the therapist's experience as well as the possibility of exploring interpersonal dynamics between the two parties. A better understanding of intra- and interpersonal dynamics requires sensitive, continuous, and objectively codified affect data (Butler, 2015).

Adar Paz https://orcid.org/0000-0001-9023-8167
Eshkol Rafaeli https://orcid.org/0000-0002-1226-9678
Sharon Gannot https://orcid.org/0000-0002-2885-170X
Bracha Laufer-Goldshtein https://orcid.org/0000-0003-2606-4793
Shrikanth Narayanan https://orcid.org/0000-0002-1052-6204
Joseph Keshet https://orcid.org/0000-0003-2332-5783

Correspondence concerning this article should be addressed to Adar Paz, Bar-Ilan University, Psychology Department, Ramat-Gan, Israel. Email: adar.paz@gmail.com

One rich source of such data, which can be collected unobtrusively, is the human voice. In this study, we examined whether the analysis of vocal features could tap into both intrapersonal and interpersonal affective arousal dynamics, which may reflect two different pathways to affect regulation within psychotherapy. Specifically, we examined whether certain intra- and interpersonal dynamic patterns facilitate favorable outcomes, both within sessions and across treatment.

Intrapersonal Affect Dynamics

Within the broader field of affect and affective regulation research, considerable attention has been paid to emotional *self*-regulation, which is defined as the activation of the goal to influence emotional trajectories (Gross, 2015), often toward re-establishing some homeostatic set point (Cole et al., 2004; Gross, 1999). Clinical theorists (e.g., Aron & Harris, 2014; Fosha, 2001; McCullough, 2003) and researchers (e.g., Greenberg, 2012) have paid growing attention to the regulation of affect as a potentially unifying target for intervention. Indeed, increases in clients' regulation capabilities have been found to predict improved outcomes (e.g., Berking et al., 2008; Pos et al., 2017; Radkovsky et al., 2014).

To date, most studies examining affect regulation within psychotherapy have relied on clients' subjective reports of their regulation skills; i.e., the explicit aspects of emotion regulation (see Gratz & Roemer, 2004; Radkovsky et al., 2014; Sloan & Kring, 2007). Self-reported measures draw heavily on clients' capability and willingness to communicate their skills and difficulties (Cummins et al., 2015). Importantly, a recent review of the literature emphasized that any understanding of intrapersonal emotion regulation would be incomplete without considering the implicit aspects of this process (Joormann & Stanton, 2016).

To go beyond self-reports, several studies have made use of observer ratings of clients' emotional or affective arousal during treatment (e.g., Carryer & Greenberg, 2010). These typically analyze data from a predefined number of sessions per treatment that are often chosen to reflect the early, middle, and late phases of therapy. For example, a recent study of clinical coders' assessments of six sessions found that emotional arousal increased across treatment and that this increase was associated with better treatment outcomes (Fisher et al., 2019; Pos et al., 2017).

Studies relying on observer coding can provide a rich and detailed view of affective arousal processes but are time consuming to conduct and limited in their ability to examine entire courses of therapy. In particular, they do not provide a way to examine moment-by-moment affect dynamics. Moreover, they are not suited for detecting or modeling characteristic high-resolution patterns such as oscillations around a set point (Boker & Nesselroade, 2002; Helm et al., 2012; Reed et al., 2015).

Another limitation of studies on affect dynamics in psychotherapy is their almost exclusive focus on clients' emotional processes (and, at times, on therapist interventions that drive them; for exceptions, see Atzil-Slonim et al., 2018; Duan & Kivlighan, 2002). Yet most affect dynamics occur in an interpersonal context, and psychotherapy is quintessentially such a context.

Interpersonal Affect Dynamics

The dyadic view of affect dynamics has been gaining increased attention in recent years (Helm et al., 2012; Reed et al., 2015) and has led to the development of considerable research on interpersonal emotion regulation (for review see Dixon-Gordon et al., 2015) or coregulation (Butler & Randall, 2013). *Interpersonal emotion regulation* encompasses regulatory strategies in which individuals use interpersonal situations to regulate their own or another's emotions (Zaki & Williams, 2013). *Coregulation* is defined as a bidirectional linkage between dyad members' oscillating emotions, which ultimately contributes to achieving an optimal level of experienced emotion in both participants (Butler & Randall, 2013). Both concepts are extremely pertinent to clinical theories (e.g., Aron & Harris, 2014; Fosha, 2001; McCullough, 2003) whose foundational principle is that the affective dyadic dynamics occurring between therapists and clients constitute key transformational agents of change in psychotherapy.

The importance of both intra- and interpersonal affect dynamics was first acknowledged by contemporary theories of affect outside of psychotherapy (e.g., developmental psychology: Feldman (2012, 2015); relationship science: Helm et al. (2012), Butler & Randall (2013), Zaki & Williams (2013)). As developmental data have shown, optimal affective states are often coconstructed dyadically during interactions in which a sensitive and responsive adult helps an infant observe and internalize regulatory skills. "Good enough" relationships are characterized by synchrony and attunement between the caregiver and the child, but this synchrony must be "marked" (Fonagy & Luyten, 2009); i.e., through attunement, the caregiver must be able to do more than simply mirror the child, but rather must soothe him or her to enable the return to an optimal arousal level. Ultimately, this dyadic affective process is expected to be internalized, as the child gradually gains the ability to provide him or herself with the same regulation initially acquired primarily through this relationship (Feldman, 2012; Ham & Tronick, 2009).

Many psychodynamic theories have highlighted the importance of client–therapist affect dynamics as promoters of clients' intrapersonal regulation abilities (e.g., Bromberg, 2003; Fosha, 2001; McCullough, 2003; Mitchell, 1993; Summers & Barber, 2009; Winnicott, 1971). When clients experience an emotion or share it with their therapist, the latter naturally reacts emotionally and the dyad's emotional responses become inextricably linked. The emotional "dance" during the client–therapist interaction, which involves a delicate balance between synchrony and discrepancy in the client's and therapist's affective experience, is considered crucial to helping clients tolerate and regulate affective arousal that is too intense or painful for them to manage alone (Aron & Harris, 2014; Fosha, 2001). In the mutual process of interpersonal affect dynamics, the clinician can provide clients with a synchronous and attuned relationship allowing for a corrective emotional experience to occur (Castonguay & Hill, 2012). The opportunity to experience one's feelings together with an authentic and emotionally present other, who is skilled in managing intense arousal, may help the client develop (or recover) more productive affect regulation capabilities.

Although the notion that client–therapist coregulation may promote client self-regulation abilities has received the most attention from psychodynamic psychotherapists, multiple therapy approaches endorse this view (e.g., Castonguay & Hill, 2012; Greenberg, 2012). This growing acknowledgment of the importance of interpersonal affect dynamics has led psychotherapy researchers to begin addressing the role of dynamics in treatment (see Koole & Tschacher, 2016 for review).

Many studies exploring interpersonal affect dynamics have concentrated on one particular dynamic—namely, synchrony—premised on the idea that therapeutic relationships involve ongoing mutual coordination or influence between therapists and clients (Atzil-Slonim & Tschacher, 2019). Several studies using objective measures (e.g., physiology: Marci et al., 2007; Tschacher & Meier, 2020; body movement: Tschacher et al., 2014) to study client–therapist synchrony have found it to be an indicator of therapeutic success. However, other studies have found more mixed effects for synchrony (e.g., with body movement: Altmann et al., 2020; Ramseyer, 2020).

One possible explanation for these inconsistent findings with respect to affect synchrony may be the wide scope of this term. Specifically, synchrony refers to any covariation between two parties. It may reflect attunement, coregulation, and *dampening*, but may also reflect mutual escalation or *amplification* (Butler, 2015). To better understand affective dynamics, these two processes need to be differentiated. *Dampening* refers to a decrease in the amplitude of affective arousal which culminates in a return to one's homeostatic baseline, whereas *amplification* refers to an increase in amplitude and a further departure from baseline. Both dampening and amplification can occur either intrapersonally (i.e., within person) or interpersonally (i.e., between person where each party "pulls" the other party's arousal toward [dampening] or away from [amplification] his or her respective baseline, see Reed et al., 2015).

To date, few psychotherapy studies have explicitly assessed dampening vs. amplification in client and therapist affect (cf., Bryan et al., 2018; Butner et al., 2017; Soma et al., 2019). Moreover, most studies exploring interpersonal nonverbal dynamics have made use of data drawn from only one (Bryan et al., 2018; Imel et al., 2014; Marci et al., 2007; Soma et al., 2019; Tschacher et al., 2014) or two (Ramseyer & Tschacher, 2014) representative sessions.

The Vocal Channel in Psychotherapy

The vocal channel (alongside several other nonverbal channels: e.g., physiology: Kleinbub, 2017; body movement: Tschacher et al., 2014) may provide a promising gateway for examining both intrapersonal and interpersonal affect dynamics. Voice is a primary channel of emotion expression and communication (Juslin & Laukka, 2003; Schuller et al., 2011) and is thus germane to both the individual and the dyad. Voice also circumvents the need to rely on subjective measures (e.g., self-reports and clinician assessments) and can be subjected to objectively codified indices. Crucially, voice (more than other channels) lends itself easily to nonobtrusive measurement.

Starting with the pioneering ideas of Rice and Wagstaff (1967), researchers have begun using speech- and voice-related measures to study psychotherapy processes, with a striking rise in studies in recent years (e.g., Tomicic et al., 2015). Vocal measures have been found helpful in identifying subtle yet clinically relevant changes in affective states in psychotherapy (e.g., Rochman & Amir, 2013).

Several vocal features have been explored in psychotherapy studies (e.g., vocal pitch [fundamental frequency; f_0] level and variability: Yang et al., 2013; f_0 range: Breznitz, 1992; speech-rate and pause variability: Mundt et al., 2012; intensity: Alpert et al., 2001). To date, the most commonly used index in psychotherapy research is f_0. Baseline f_0 and deviations from this baseline have been shown to be strongly correlated with self-reported and physiological indicators of affective arousal (heart rate, blood pressure, and cortisol secretion; Juslin & Scherer, 2005). For example, Imel et al. (2014) showed that client–therapist vocal synchrony was linked to therapist empathy as assessed by external raters. This finding is consistent with data from nonclinical relationships associating vocal synchrony with positive relationship outcomes (see Lee et al., 2010; Lubold & Pon-Barry, 2014).

Three more recent studies have explored interpersonal affect dynamics in psychotherapy using vocal features. Gaume et al. (2019) used two large samples and attempted (but failed) to replicate the association between client–therapist vocal synchrony and therapist empathy reported by Imel et al. (2014). Bryan et al. (2018) assessed client and therapist vocal arousal (VA) during a single crisis intervention session and found that mutual dampening of affective arousal was associated with a stronger client-reported emotional bond. Finally, Soma et al. (2019) used VA measures and dynamic systems models to assess whether and how clients and therapists modulated each other's affect. They found that when clients became more emotionally labile over the course of a session, therapists became less so, and vice-versa. Furthermore, when therapists manifested greater arousal or increased in their arousal levels, clients returned to their homeostatic baseline more rapidly.

The studies noted above have all dealt with one vocal feature (f_0) and used vocal data from a single session. However, recent work on the analysis of vocal arousal suggests that a combination of several features, rather than f_0 alone, may reflect human affective arousal more accurately (Bone, Lee, & Narayanan, 2014a; Chaspari et al., 2017). In particular, an index combining intensity and pitch was found to work better than separate indices of intensity and pitch (Bone et al., 2014a, 2014b; Chaspari et al., 2017).

Furthermore, growing evidence suggests that intra- and interpersonal dynamics in vocal indicators of arousal may change as relationships unfold over time (Levitan & Hirschberg, 2011). Examining these dynamics with multifeature data collected over time could help clarify patterns of change in both client and therapist vocal arousal over the course of treatment, and thus serve to better identify intrapersonal and interpersonal affect dynamics that may predict treatment outcome.

Nevertheless, utilizing the vocal channel to examine mutual influence in affective arousal poses several challenges. Unlike other measures of arousal (e.g., electrophysiology), human conversation has a turn-taking structure, where for the most part speakers do not speak simultaneously. In addition, speech involves moments of silence (especially in psychodynamic psychotherapy). Therefore, analyses and modeling of the signal extracted from dyadic speech must contend with the fact that this signal is neither continuous nor simultaneous.

In one possible solution to these challenges, Soma et al. (2019) "smoothed" the extracted vocal features of both client and therapist over three speech turns and thus created a semi-simultaneous data series. This method has some advantages (e.g., allowing for second-order analyses) but is based on an embedded assumption that the speakers are more or less balanced in the duration of their speech turns. By contrast this study examined psychodynamic psychotherapy sessions in which speech is typically unbalanced, where clients usually speak more than therapists. For example, analyzing psychodynamic sessions, Shapira et al. (2020) found that 78% of all session utterances were made by the clients and 22% by the therapists.

Hence, we opted to use the speech analysis method based on speech-turn switches proposed by Levitan and Hirschberg (2011) who found that vocal features surrounding *turn-switches* carry more information about the emotional interaction between speakers than do average vocal scores from entire speech turns. Using this approach, and examining the parties' arousal measures in switch moments, we obtained data that were continuous without assuming that they were entirely simultaneous.

These data lend themselves to first-order dynamic systems analyses which can examine momentary changes in arousal and the direction of change (toward or away from the speakers' arousal baseline), and thus answer our key questions of whether there is a mutual influence such that one party "pulls" the other party toward their baseline (dampening) or away from it (amplification) in interpersonal dynamics, as well as whether either party's own data are marked by a return to or a departure from their baseline, in intrapersonal dynamics. Our models were based on a method proposed by Butner et al. (2018) for estimating intra (within) and inter (between) affect dynamics among romantic couples using a first-order dynamic systems model (see also Butner et al., (2017) and Chaspari et al., (2017) for additional applications of this model).

The Current Study

In line with current psychodynamic theoretical approaches (e.g., Fosha, 2001; McCullough, 2003), contemporary inter- and intrapersonal affect regulation conceptualizations (e.g., Butler, 2015) and studies indicating the association between improvement in affect regulation abilities throughout treatment and treatment outcomes (Fisher et al., 2019; Pos et al., 2017; Radkovsky et al., 2014), the following hypotheses guided our study:

1. Intrapersonal and interpersonal dampening in vocal affective arousal. We expected both clients and therapists to exhibit vocal affective dampening on average during the session; i.e., we expected their arousal level to show a "pull" toward their baseline or homeostatic set point (Hypothesis 1a). In addition, we expected the VA level of both clients and therapists to be "pulled," on average, toward their partners' arousal level, in a way that would also lead to dampening (Hypothesis 1b).

2. Treatment-level change in intrapersonal and interpersonal dampening. We expected both clients' and therapists' intrapersonal dampening in vocal arousal (i.e., the strength of the "pull" noted above) to increase over the course of treatment (Hypothesis 2a). In addition, we expected both parties' interpersonal dampening (i.e., the strength of their "pull" toward their partner's arousal level) to also increase over the course of treatment (Hypothesis 2b).

3. Association between intrapersonal and interpersonal dampening and treatment outcome. We expected clients' intrapersonal dampening in vocal arousal to be positively associated with session outcome (Hypothesis 3a). Also, we expected clients' interpersonal dampening to be positively associated with session outcome (Hypothesis 3b).

Method

Participants and Treatment

The data used in this study were obtained from the recordings of multi-session therapies conducted with 30 adult individual therapy clients at a university-based community mental health clinic. All therapies took place between August 2017 and August 2019. To be considered for inclusion in the study, the therapy needed to have included recordings from dual microphones and to have lasted at least 15 sessions; of these, at least every other session needed to have adequate audio quality as well as both pre- and post-session Outcome Rating Scale (ORS; Miller et al., 2003) measurements. The sample consisted of 279 therapy sessions from 30 clients treated by 24 therapists, with a mean of $M = 9.3$ ($SD = 2.41$) sessions per dyad. The average number of days between consecutive session was $M = 8.43$ ($SD = 3.39$) with a mode and a median of 7 and a range of 3–29.

Clients

The clients included in the sample received an average of 26.4 treatment sessions ($SD = 5.33$, range [15,36]). Their mean age was 35.6 years ($SD = 12.5$, range [21,69] years). The majority of the clients were female (66.6%). The Mini-International Neuropsychiatric Interview version 5.0 (M.I.N.I; Sheehan et al., 1998) was used to establish Axis I diagnoses for these clients. Of the total sample, 40% had a single diagnosis, 15% had two diagnoses, and 21% had three or more diagnoses. Most clients were diagnosed with affective disorders (43%) or anxiety disorders (23%) as the primary diagnosis. Additional primary diagnoses included obsessive–compulsive disorder (4%) or other disorders (7%).

Therapists

Twenty-four therapists were included in the sample (77% female). Twenty-two therapists were MA students with a range of 0–30 previous clinical hours; two therapists were PhD students with a range of 50–250 previous clinical hours. Eighteen therapists treated one client and six therapists treated two clients. The therapists received 1 hr of individual supervision and 4 hr of group supervision on a weekly basis. The supervisors were senior clinicians with expertise in psychodynamic models.

Treatments

Individual psychotherapy consisted of 1–2 weekly sessions. The dominant approach in the clinic is a short-term psychodynamic psychotherapy treatment model based on a blend of object relations, self-psychology, and relational theories (Kohut, 1971; Winnicott, 1971). The key features of the model included (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, and fantasies (e.g., Shedler, 2010; Summers & Barber, 2009). Treatment was open-ended in length; however, given that it was conducted at a university-based clinic following an academic calendar, treatments lasted between 9 months to 1 year.

Measures

Outcome Rating Scale

The ORS (Miller et al., 2003) is a four-item visual analog scale developed as a brief alternative to longer outcome measures. The scale is designed to assess change in three areas of client functioning that are widely considered as valid indicators of progress in treatment: functioning, interpersonal relationships, and social role performance. Respondents complete the ORS by rating the items on a visual analog scale anchored at one end by the word Low and at the other end by the word High. The sum of the items ranges from 0 to 40, with higher scores indicating better functioning. The ORS was completed twice each session: immediately before and after the session. Subsequently, pre-to-post ORS change (ORS_diff) was calculated as the pre-session ORS subtracted from the post-session ORS.

Vocal Arousal (VA)

A multifeature vocal arousal extraction tool was used (Bone et al., 2014a, 2014b). The original audio was segmented into speech turns, using an automatic diarization algorithm developed specifically for psychotherapy conversations, as speakers in such conversations typically present unbalanced activity patterns. Specifically, clients often speak for longer periods, while therapists frequently respond with shorter utterances. To address this, we used an algorithm based on previous work on speech diarization and separation (Laufer-Goldshtein et al., 2018a, 2018b) whose proprietary method has been submitted for publication (Laufer-Goldshtein et al., Submitted).[1]

Subsequently, following Bone et al. (2014a), VA was computed as a weighted average index of three speech features: (a) intensity, (b) pitch, and (c) HF500 (the ratio of energy above 500 Hz divided by the energy between 80 Hz and 500 Hz). These features were then normalized for each participant for each session allowing for the average level of each feature to act as the speaker's "baseline." The final VA score was created from the weighted average of the three feature scores. This measure has achieved state-of-the-art performance for cross-corpus automatic arousal recognition (Valstar et al., 2016). Across the 30 therapy dyads and 279 available sessions, there were 21,133 mean VA observations ($M = 75.5$ observations per session [$SD = 41.9$]).[2]

Procedure

The study was conducted in compliance with the University's ethical review board. The data were obtained as part of the routine monitoring used in the clinic. Clients consented to participate voluntarily and were told that they could terminate their participation at any time with no effect on their treatment and that the therapists would be unaware of their responses. The clients completed the ORS electronically (using computers located in the clinic rooms) before and after each therapy session.

Data Analysis

In the analysis, we followed the approach described in Levitan and Hirschberg (2011) who found that vocal features surrounding *turn-switches* carry more information about the affective interaction between the speakers than do average vocal scores from entire speech turns. For this purpose, they suggested focusing on interpausal units (IPUs); i.e., parts of speech-turns that are demarcated by pauses lasting at least 50 ms, and which themselves are pause-free (i.e., interrupted, at most, by pauses lasting less than 50 ms).

Accordingly, to capture the interpersonal affect dynamics unfolding between speakers, we defined the basic unit of analysis as dyads' turn-switches; i.e., the last IPU in Speaker A's speech turn followed by the first IPU in Speaker B's subsequent speech turn (see Figure 1).

The models were based on a method proposed by Butner et al. (2018, 2017) for estimating dyadic affect dynamics in romantic couples using a first-order dynamic systems model. Specifically, in this model, the first derivative (i.e., the first-order change in Partner A's VA from speech turn $i-1$ to i) is predicted by Partner A's and Partner B's previous VA assessments. Since the data were nested (speech turn switches nested within sessions, which themselves were nested within dyads), we used a multivariate multilevel framework (Baldwin et al., 2014). In this framework, the first derivatives of the clients' and therapists' VA levels were modeled simultaneously, whereas their residuals were allowed to vary within session (Level 1), between session (Level 2), and between dyads (Level 3).[3]

In the following section, the three models underlying the study hypotheses are presented.

Model 1 [Hypotheses 1a & 1b]: Average Intra- and Interpersonal Dampening

$$\left(\frac{\mathrm{VA}^{c}_{(2i)sd}-\mathrm{VA}^{c}_{(2i-1)sd}}{\Delta t}\right)=\gamma^{c}_{000}+\gamma^{c}_{100}*\mathrm{VA}^{c}_{(2i-1)sd}+\gamma^{c}_{200}*\left(\frac{\mathrm{VA}^{c}_{(2i-1)sd}-\mathrm{VA}^{t}_{(2i)sd}}{\Delta t}\right)+u^{c}_{ood}+r^{c}_{0sd}+e^{c}_{(2i)sd} \quad (1)$$

$$\left(\frac{\mathrm{VA}^{t}_{(2i+1)sd}-\mathrm{VA}^{t}_{(2i)sd}}{\Delta t}\right)=\gamma^{t}_{000}+\gamma^{t}_{100}*\mathrm{VA}^{t}_{(2i)sd}+\gamma^{t}_{200}*\left(\frac{\mathrm{VA}^{t}_{(2i)sd}-\mathrm{VA}^{c}_{(2i+1)sd}}{\Delta t}\right)+u^{t}_{ood}+r^{t}_{0sd}+e^{t}_{(2i+1)sd} \quad (2)$$

Equations 1 & 2: Dynamics systems multilevel model of clients' (Equation 1) and therapists' (Equation 2) inter- and intra-VA affect dynamics.

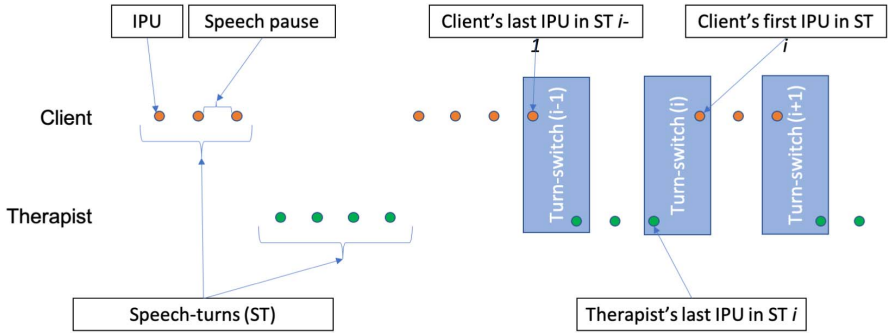
In this model, VA change (i.e., the first derivative) of client c (or therapist t) in IPU $2i$ in session s in client–therapist dyad d is predicted by this client's (or therapist's) intercept (γ^{c}_{000} *or* γ^{t}_{000}). It is also predicted by this client's (or therapist's) previous IPU's VA (γ^{c}_{100} *or* γ^{t}_{100}). Notably, when this parameter (γ^{c}_{100} *or* γ^{t}_{100}) is negative, it indicates that the speaker's VA is "attracted" to their baseline; i.e., when in one IPU the speaker has negative (i.e., below baseline) VA, they will tend to show an increase (i.e., a positive first

[1] Additional information and details on the diarization method can be found in the Online Supplementary Materials (OSM): https://osf.io/hrgcb/

[2] The diarization algorithm and VA extraction implemented MATLAB (Version 2019a). The vocal features were extracted using Praat (Boersma & Weenink, 2017). For additional information, see https://osf.io/hrgcb/).

[3] Because six therapists treated two patients each, we also tested whether nesting affected the results; it did not. For more information, please see OSM (https://osf.io/hrgcb/)

Figure 1
Speech-Turns, Pauses, IPUs, and Turn-Switches



Note. See the online article for the color version of this figure.

derivative) in their VA in the following IPU, and vice versa (which makes this parameter relevant to Hypothesis 1a). In addition, the speaker's VA change is predicted by the difference between this speaker's VA in IPU $i-1$ and the partner's VA in IPU i (γ^{c}_{200} *or* γ^{t}_{200}). The difference between the two parties' VA is scaled by the length of time between these two IPUs ($\frac{1}{\Delta t}$), to account for variation in the duration of speech turns. Note that when this parameter (γ^{c}_{200} *or* γ^{t}_{200}) is negative, it indicates that the speaker's VA is "attracted" toward the partner's VA—i.e., when in one speech-turn (e.g., speech turn $i-1$) the two parties' VA difference is negative (i.e., one party's VA is above the other's), the speaker will tend to show an increase (i.e., a positive first derivative) in their VA in the following IPU, and vice versa (which makes this parameter relevant to Hypothesis 1b). Finally, to account for data nesting, the model includes random effects at the level of the dyad (u^{c}_{ood} *or* u^{t}_{ood}), the session (r^{c}_{osd} *or* r^{t}_{osd}), and the IPU (e^{c}_{isd} *or* e^{t}_{isd}). Notably, these two equations were run simultaneously to allow clients' and therapists' Level-2 and Level-3 residual terms to covary (thus accounting for the dyads' interdependence).

Model 2 [Hypotheses 2a & 2b]: Change in Intra- and Interpersonal Dampening

The second model was designed to test whether intrapersonal and interpersonal dampening in VA increased during therapy. To do so, we added the main effect of session number[4], as well as its interaction with both intrapersonal and interpersonal dampening, to Model 1.

Model 3 [Hypotheses 3a and 3b]: Intra- and Interpersonal Dampening and Session Outcome

The third model was designed to test whether and inter-personal dampening were stronger in good outcome sessions. This was done by testing whether the session outcome (a Level-2 variable) moderated the two dampening parameters. Session outcome was operationalized as the difference between the client's well-being reported at the end vs. the beginning of sessions. We added the main effect of this difference score as well as its cross-level interaction with intrapersonal and interpersonal dampening to Model 1.

Results

Model 1: Average Intra- and Interpersonal Dampening

Table 1 presents the fixed effects estimated in Model 1. As can be seen, and in line with Hypothesis 1a, the effect pertaining to intrapersonal dampening was negative and significant for both clients and therapists. In other words, both clients' and therapists' VA levels were "pulled" toward their own baseline. Additionally, this "pull" was stronger the more the VA scores (either therapists' or clients') deviated from their own baseline. Consistent with Hypothesis 1b, the effect pertaining to "interpersonal dampening" was negative and significant for both clients and therapists. In other words, both clients' and therapists' VA levels were also "pulled" toward the other party's VA. Additionally, this "pull" was stronger the greater the difference in VA levels between the two parties.

Notably, a post hoc analysis revealed that the clients' intrapersonal and interpersonal dampening parameters were significantly stronger than those of the therapists (intrapersonal: Est. = −.08, SE = .01, $p < .001$; interpersonal: Est. = −.10, SE = .01, $p < .001$; see Figure 2).

Theoretically, one could argue that this interpersonal "pull" does not necessarily entail interpersonal dampening, since one can also be "pulled" by their partner away from their baseline. For interpersonal dampening to occur (when the interpersonal "pull" is toward the baseline), the partner's VA should be positioned closer to the baseline than the speaker's VA. To test this, we ran an additional 3-level ML analysis in which the difference between one's VA in speech turn $i-1$ and one's partner's VA in speech turn i was predicted by one's VA in speech turn $i-1$. In this analysis,[5] a positive and significant estimate emerged for both clients (Est. = .41, SE = .02, $p < .001$) and therapists (Est. = .36, SE = .02, $p < .001$), indicating that when one party's VA deviated from his or her own baseline, the other party's VA tended to be positioned

[4] To address one reviewer's concern, we re-ran the analyses with elapsed days in the therapy instead of session number. The results remained similar with minor exceptions. For more information, see OSM (https://osf.io/hrgcb/).

[5] The interpersonal prediction in Equations 1 & 2 (Model 1) allowed us to estimate the extent to which one speaker's affective arousal impacted arousal change in the other speaker. Distinguishing between dampening and amplification required an additional analysis; see the OSM (https://osf.io/hrgcb/) for more information.

Table 1
Fixed Effect Predictors for Speakers' VA Change

	Client			Therapist		
	Est.(*SE*)	CI(95%)	*p*	Est.(*SE*)	CI(95%)	*p*
Intercept ($\gamma_{000}^{c\|t}$)	.037(.02)	[−.002, .075]	.06	−.036(.01)	[−.057, −.015]	.001
Intrapersonal dampening ($\gamma_{100}^{c\|t}$)	−.412(.01)	[−.432, −.393]	<.001	−.333(.01)	[−.349, −.317]	<.001
Interpersonal dampening ($\gamma_{200}^{c\|t}$)	−.323(.01)	[−.344, −.303]	<.001	−.219(.01)	[−.234, −.204]	<.001

closer to the baseline. Together with the finding that partners' VA levels served as "attractors" for each other, it indicates that dyad members tended to exhibit interpersonal dampening (rather than amplification).

Model 2: Change in Intra- and Interpersonal Dampening Throughout Treatment

To test whether intra- and interpersonal dampening increased over the course of treatment, the main effect of session number and interaction with these two terms were added to Model 1. As shown in Table 2, there was a significant interaction between session number and clients' intrapersonal dampening. Thus, in line with Hypothesis 2a, clients' intrapersonal dampening increased as the treatment progressed. No such interaction was found for therapists. In addition, a significant interaction between session number and clients' interpersonal dampening was found. However, in contrast to Hypothesis 2b, clients' interpersonal dampening actually decreased as the treatment progressed. No such interaction was found for therapists. Figure 3 illustrates the increase in intrapersonal dampening and decrease in interpersonal dampening found among clients over the course of therapy.

Model 3: Intra- and Interpersonal Dampening and Session Outcome

To test the association between session outcome on the one hand and intra- and interpersonal dampening on the other, the main effect of session outcome as well as its interaction with the two terms was added to Model 1. As shown in Table 3, disconfirming Hypothesis 3a, no interaction was found between session outcome and clients' intrapersonal dampening. However, there was a significant interaction between session outcome and clients' interpersonal dampening. To probe this interaction, we computed the parameters at 1 *SD* above and below the baseline. When session outcomes were poor (i.e., 1 *SD* below the mean), the parameter was weaker (Est. = −.29, *SE* = .01, $p < .001$) than when session outcomes were good (i.e., 1 *SD* above the mean; Est. = −.37, *SE* = .01, $p < .001$). Thus, in line with Hypothesis 3b, clients' interpersonal dampening was positively associated with session outcome. No interaction effects were found for the therapists' intra- or interpersonal dampening.[6] Figure 4 illustrates the associations between intrapersonal dampening and session outcome as well as interpersonal dampening and session outcome.

Discussion

We used vocal measures and dynamic system models to examine intrapersonal and interpersonal affect dynamics within and between clients and therapists; we also examined the development of these dynamics across treatments and their associations with treatment outcome. Consistent with contemporary emotion regulation conceptualizations (Butler & Randall, 2013; Thompson, 2011; Zaki & Williams, 2013), we found both within-person (intrapersonal) and between-person (interpersonal) affect regulatory dynamics.

As predicted, both clients' and therapists' VA evidenced intrapersonal dampening (Hypothesis 1a). Specifically, although both therapists and clients departed from their arousal baseline, their VA levels were "pulled" back to these baselines. This pattern is consistent with previous intrapersonal emotion regulation results reported in clinical settings (Soma et al., 2019) and also with findings for romantic couple data (including both self-reported emotions [Butner et al., 2007] and heart rate [Helm et al., 2012]).

Consistent with our next prediction, both clients and therapists exhibited interpersonal affect dampening (Hypothesis 1b). Specifically, both the clients' and the therapists' levels of arousal were "pulled" toward the other party's arousal level. Additionally, therapists' arousal levels were closer to baseline (on average) than their clients. Taken together, these findings suggest that on average, clients were "pulled" by their therapists' VA toward their own baseline.

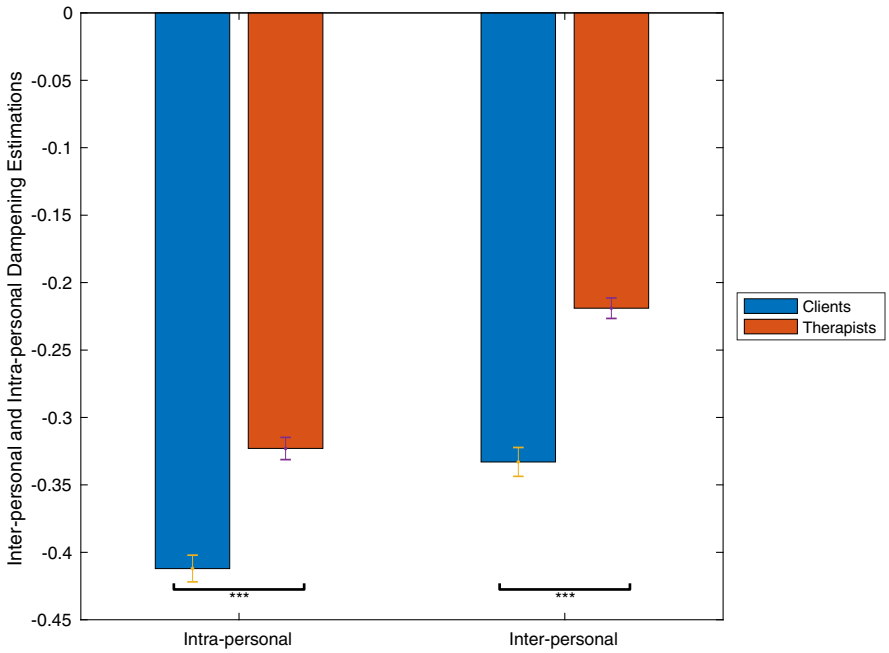
Our results suggest that intrapersonal and interpersonal dynamics occur simultaneously, which may suggest that both internal and external resources are used for affective arousal regulation (Dixon-Gordon et al., 2015; Uchino et al., 1996; Zaki & Williams, 2013). This type of pattern is consistent with previous findings (e.g., Soma et al., 2019). However, whereas Soma and her colleagues (as well as Bryan et al., 2018) used single sessions to assess these dynamics and relied on vocal pitch as their key measure of affective arousal, we examined these processes session-by-session throughout treatment and implemented recent advances in signal processing by using multiple vocal features (Bone et al., 2014a).

When we examined the pattern of change in intrapersonal affect dampening over the course of treatment, we found that it rose as therapy progressed (Hypothesis 2a). However, contrary to our hypothesis (Hypothesis 2b), we found that interpersonal dampening actually decreased throughout treatment. These findings echo recent psychodynamic psychotherapy theories which suggest that throughout treatment, clients expand their affect regulation capacities resulting from the combined resources of the dyad which they eventually internalize and "make their own" (Aron & Harris, 2014; Fosha, 2001). It is possible that throughout treatment as client self-regulation capacity increases, the need for the therapist as an external source of regulation gradually decreases (Summers & Barber, 2009). In the current study, we examined this theoretical

[6] To examine the simple effects of the clients' interpersonal dampening at various levels of session outcome (i.e., low [−1 *SD*], average, and high [+1 *SD*] ORS difference), we used Preacher et al. (2006) computational tool for probing interaction effects in MLM analyses.

Figure 2

Model 1: Intra- and Interpersonal Dampening Coefficients Contrasts Between Clients and Therapists



Note. See the online article for the color version of this figure.

idea in a sample of clients in psychodynamic psychotherapy. However, the idea that the client–therapist emotional experience in psychotherapy sessions promotes clients' abilities to regulate their affect is central to many psychotherapy approaches (e.g., Greenberg, 2012; Rafaeli et al., 2010). Soma et al., (2019) which have reported client–therapist coregulation processes in Motivational Interviewing. Future studies should examine these processes with clients treated by other forms of psychotherapy.

Finally, our third hypothesis was only partially supported by the results. Although contrary to expectations (Hypothesis 3a), we did not find an association between clients' intrapersonal dampening and session outcome, our results were in line with Hypothesis 3b. Specifically, sessions in which the clients manifested inter-personal dampening were also the ones in which they evidenced a larger reduction in symptoms, from pre- to post-session. It is possible that within-session gains (improvement in well-being from pre- to post-session) are more highly affected by interpersonal processes between the client and the therapist than by intrapersonal processes that occur within the client. The association between interpersonal arousal dampening and session outcome is consistent with both social baseline and attachment theories, which predict that people often rely on each other as regulatory resources to enhance their feeling of security (Beckes & Coan, 2011; Sbarra & Hazan, 2008). These results are also consistent with several psychodynamic models (e.g., Benjamin, 2003; Fonagy et al., 2018; Fosha, 2001; Winnicott, 1971) which highlight the importance of emotional connection (previously found to be related to greater empathy; Bryan et al., 2018; Imel et al., 2014; cf., Gaume et al., 2019).

This study extends earlier work exploring moment-to-moment affect dynamics among clients and therapists (Bryan et al., 2018; Soma et al., 2019) in several ways. First, most previous studies of interpersonal dynamics have focused on client–therapist synchrony (Altmann et al., 2020; Kleinbub, 2017; Marci et al., 2007; Ramseyer & Tshcacher, 2014). Our study examined interpersonal

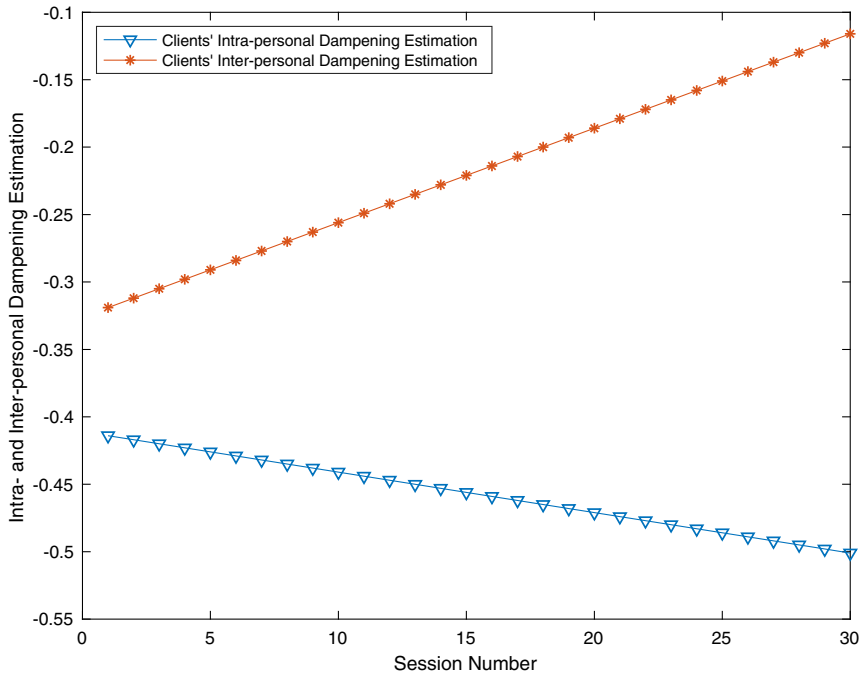
Table 2

Fixed Effect Predictors for the Modeled Interaction Between Interpersonal Dampening, Intrapersonal Dampening, and Session Number

	Client			Therapist		
	Est.(*SE*)	CI(95%)	*p*	Est.(*SE*)	CI(95%)	*p*
Intercept	.036(.02)	[−.001, .074]	.059	−.036(.01)	[−.057, −.014]	.001
Intrapersonal dampening	−.411(.01)	[−.431, −.392]	<.001	−.333(.01)	[−.349, −.317]	<.001
Interpersonal dampening	−.326(.01)	[−.347, −.305]	<.001	−.22(.01)	[−.235, −.205]	<.001
Session number	−.002(.001)	[−.005, .001]	.175	.001(.001)	[−.001, .004]	.297
Intrapersonal dampening × Session number	−.003(.001)	[−.007, −.001]	.034	−.001(.001)	[−.003, .002]	.678
Interpersonal dampening × Session number	.007(.001)	[.004, .01]	<.001	.001(.001)	[−.001, .004]	.323

Figure 3

Clients' Intrapersonal and Interpersonal Dampening Over the Course of Therapy



Note. See the online article for the color version of this figure.

dampening (or what Butler and Randall (2013) refer to as *coregulation*); i.e., a specific form of synchrony in which one person's affect influences their partner's affect to return the homeostasis baseline. Second, this is the first project to assess intra- and interpersonal affect dynamics over the course of multi-session treatment. Finally, this study used a multifeature index of vocal arousal (Bone et al., 2014a) as well as a novel model to simultaneously assess intra- and interpersonal dynamics for both clients and therapists.

Limitations, Future Directions, and Summary

This study's contributions should be considered in light of its limitations. One limitation was the relatively small sample of dyads (i.e., 30 dyads). It may be that this study was underpowered to detect between-dyad effects as well as smaller between-session effects, such as the hypothesized associations between therapist affect dynamics and session outcome, or the hypothesized change in the therapists' affect dynamic patterns over the course of treatment. As such, our findings may be regarded as preliminary until replicated, though they now provide a good starting point for future *a priori* power analyses.

Our decision to focus on turn switches has not only some benefits (outlined earlier) but also some costs. Key among these is the fact that much of the VA data (occurring outside of the switches) were excluded from analysis. We considered the remaining data as the most appropriate for modeling first-order but not second-order dynamics (see Butler et al., 2017). First-order dynamics can inform us about the direction and level of VA change; i.e., whether VA has shifted toward or away from baseline. In contrast, more continuous VA data would lend themselves to the modeling of second-order dynamics, which would account for the *rates* at which the speakers' VA levels change. It would be interesting, for example, to examine how quickly particular clients dampen (or amplify) their arousal levels, or whether this rate of change itself changes as therapy

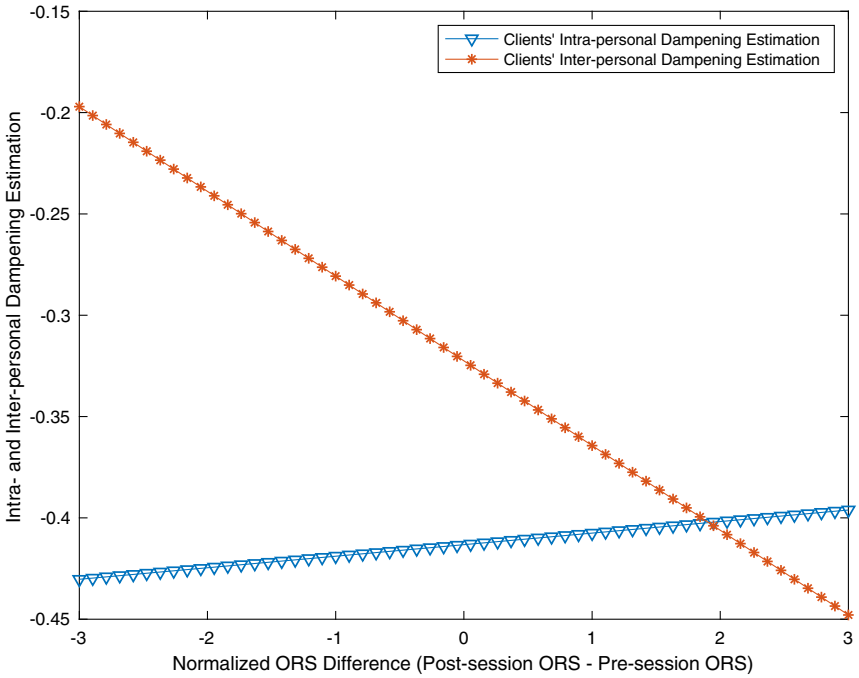
Table 3

Fixed Effect Predictors for the Modeled Interaction Between Interpersonal Dampening and Session Outcome

	Client			Therapist		
	Est.(*SE*)	CI(95%)	*p*	Est.(*SE*)	CI(95%)	*p*
Intercept	.037(.02)	[−.001, .076]	.057	−.036(.01)	[−.057, −.014]	.001
Intrapersonal dampening	−.413(.01)	[−.432, −.393]	<.001	−.333(.01)	[−.349, −.317]	<.001
Interpersonal dampening	−.324(.01)	[−.345, −.303]	<.001	−.219(.01)	[−.234, −.205]	<.001
ORS diff.	.006(.01)	[−.012, .024]	.553	−.009(.01)	[−.025, .006]	.238
Intrapersonal dampening × ORS diff.	.006(.01)	[−.014, .027]	.551	−.012(.01)	[−.03, .006]	.180
Interpersonal dampening × ORS diff.	−.044(.01)	[−.065, −.023]	<.001	.012(.01)	[−.004,.028]	.141

Figure 4

The Association Between Session Outcome and Clients' Intrapersonal and Interpersonal Dampening



Note. See the online article for the color version of this figure.

progresses. As noted above, such analyses form one of the advantages of the approach used by Soma et al. (2019).

The framework for this work was that of dynamic systems modeling; as such, caution should be exercised when interpreting each coefficient by itself. There can be higher order effects (i.e., second-order effect; see Butler et al., 2017) but also an interaction between the effects found individually and between the parties. Future studies could examine the association between intrapersonal dampening levels and the interpersonal as well as between speakers.

We found that clients' interpersonal dampening was generally associated with better outcomes. However, for some clients and/or some sessions, amplification may have been even more therapeutic. This idea is inherent to clinical theories that emphasize emotion activation (e.g., Carryer & Greenberg, 2010). Future studies should thus explore the possibility that interpersonal emotional amplification may play an adaptive role in successful psychotherapy under certain conditions (e.g., depressed people who tend to experience more blunted emotions and lower arousal levels [e.g., Bylsma et al., 2008] and who may therefore benefit from activation rather than dampening).

Relatedly, we focused our attention on affective arousal but did not distinguish between different types of emotions (e.g., sadness, anger, happiness). Moreover, though we attended to one central dimension of affective space (i.e., arousal), we did not consider its interaction with the other central dimension defining this space; i.e., valence (Tellegen et al., 1999). Clearly, a more fine-grained account of the specific emotion or at least the valence of the emotion present would be very informative. This is because some therapy sessions (or even entire courses of treatment) may be aimed less at downregulation (e.g., of distress) and more at upregulation (e.g., of assertive anger, playfulness, or expression of suppressed wishes). To further our understanding of the intra- and interpersonal affect dynamics that may accompany such cases, we need more information about the specific emotions (including their valence), as well as the client's specific emotional goals.

At present, our data point to overall dampening vs. amplification patterns, but at this stage, tell us little about the context, topics, or specific interventions that accompanied or were addressed within the sessions. For example, in the current sample (of clients undergoing psychodynamic psychotherapy), therapists might have used questions, reflections, interpretations, or confrontations, and these may have had differential effects on the clients' arousal levels (at times dampening them and at other times amplifying them). The methods used in the present study could prove useful for future naturalistic research investigating such questions. This research could examine which conditions, therapist stances, or treatment interventions facilitate affect dampening vs. amplification.

Taken together, our results highlight the potential of computerized vocal analyses as a way of looking at moment-by-moment processes within psychotherapy sessions. They point to an interesting pattern of results related to affect dampening in which the interpersonal route (where therapists' affective arousal "pulls" their clients toward homeostasis) appears to play as large a role, if not larger, than the intrapersonal route (where clients' affect arousal is pulled toward its own baseline). Though the findings are only preliminary, they may have clinical implications in terms of

elucidating the therapist's role in helping clients who experience high affective arousal in therapy sessions. The findings may assist therapists in finding a proper voice with their clients, by suggesting that it is important to be in sync with the client's fluctuating affective arousal, but also to regulate one's own arousal to be better able to help clients downregulate their painful emotions, which may in turn lead to better therapy outcomes.

Data Transparency

The data reported in this manuscript were previously published as part of a larger data collection, continuously collected in the Bar-Ilan psychology department's outpatient clinic. Findings from the data collection have been reported in separate manuscripts. Fisher et al. (2019) focus on self-reported "emotional-experience" and "self-understanding" and the association with clients' functioning; Bar-Kalifa & Atzil-Slonim (2020) focuses on intrapersonal and client-therapist interpersonal emotional dynamics as measured by self-reports session-to-session. Both these previous studies utilized self-reports of clients' emotions, whereas in the current project, we used speech recordings that were not examined in these studies. The current paper deals with intra- and interpersonal emotional dynamics in high-resolution time-series moment-by-moment within sessions and the associations between these dynamics to session outcomes.

References

Alpert, M., Pouget, E. R., & Silva, R. R. (2001). Reflections of depression in acoustic measures of the patient's speech. *Journal of Affective Disorders*, *66*(1), 59–69. https://doi.org/10.1016/S0165-0327(00)00335-9

Altmann, U., Schoenherr, D., Paulick, J., Deisenhofer, A. K., Schwartz, B., Rubel, J. A., Stangier, U., Lutz, W., & Strauss, B. (2020). Associations between movement synchrony and outcome in patients with social anxiety disorder: Evidence for treatment specific effects. *Psychotherapy Research*, *30*(5), 574–590.

Aron, L., & Harris, A. (2014). Relational psychoanalysis: Vol. 5. *Evolution of process*. Routledge.

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*(1), 51–64. https://doi.org/10.1037/cou0000250

Atzil-Slonim, D., & Tschacher, W. (2019). Dynamic dyadic processes in psychotherapy: Introduction to a special section. *Psychotherapy Research*, *30*, 555–557. https://doi.org/10.1080/10503307.2019.1662509

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*(5), 920–930. https://doi.org/10.1037/a0035628

Bar-Kalifa, E., & Atzil-Slonim, D. (2020). Intrapersonal and interpersonal emotional networks and their associations with treatment outcome. *Journal of counseling psychology*, *67*(5). https://doi.org/10.1037/cou0000415

Beckes, L., & Coan, J. A. (2011). Social baseline theory: The role of social proximity in emotion and economy of action. *Social and Personality Psychology Compass*, *5*(12), 976–988. https://doi.org/10.1111/j.1751-9004.2011.00400.x

Benjamin, L. S. (2003). *Interpersonal reconstructive therapy: Promoting change in nonresponders*. The Guilford Press.

Berking, M., Wupperman, P., Reichardt, A., Pejic, T., Dippel, A., & Znoj, H. (2008). Emotion-Regulation skills as a treatment target in psychotherapy. *Behaviour Research and Therapy*, *46*(11), 1230–1237. https://doi.org/10.1016/j.brat.2008.08.005

Boersma, P., & Weenink, D. (2017). *Praat: Doing phonetics by computer* (Version 6.0.24) [Computer program]. http://www.praat.org/

Boker, S. M., & Nesselroade, J. R. (2002). A method for modeling the intrinsic dynamics of intraindividual variability: Recovering the parameters of simulated oscillators in multi-wave panel data. *Multivariate Behavioral Research*, *37*(1), 127–160. https://doi.org/10.1207/S15327906MBR3701_06

Bone, D., Lee, C. C., & Narayanan, S. (2014a). Robust unsupervised arousal rating: A rule-based framework with knowledge-inspired vocal features. *IEEE Transactions on Affective Computing*, *5*(2), 201–213. https://doi.org/10.1109/TAFFC.2014.2326393

Bone, D., Lee, C. C., Potamianos, A., & Narayanan, S. S. (2014b). *An investigation of vocal arousal dynamics in child-psychologist interactions using synchrony measures and a conversation-based model*. Fifteenth Annual Conference of the International Speech Communication Association, Singapore, September 14–18, 2014.

Breznitz, Z. (1992). Verbal indicators of depression. *Journal of General Psychology*, *119*, 351–363. https://doi.org/10.1080/00221309.1992.9921178

Bromberg, P. M. (2003). Something wicked this way comes: Trauma, dissociation, and conflict: The space where psychoanalysis, cognitive science, and neuroscience overlap. *Psychoanalytic Psychology*, *20*(3), 558–574. https://doi.org/10.1037/0736-9735.20.3.558

Bryan, C. J., Baucom, B. R., Crenshaw, A. O., Imel, Z., Atkins, D. C., Clemans, T. A., Leeson, B., Burch, T. S., Mintz, J., & Rudd, M. D. (2018). Associations of patient-rated emotional bond and vocally encoded emotional arousal among clinicians and acutely suicidal military personnel. *Journal of Consulting and Clinical Psychology*, *86*(4), 372–383. https://doi.org/10.1037/ccp0000295

Butler, E. A. (2015). Interpersonal affect dynamics: It takes two (and time) to tango. *Emotion Review*, *7*(4), 336–341. https://doi.org/10.1177/1754073915590622

Butler, E. A. (2017). Emotions are temporal interpersonal systems. *Current Opinion in Psychology*, *17*, 129–134. https://doi.org/10.1016/j.copsyc.2017.07.005

Butler, E. A., & Randall, A. K. (2013). Emotional coregulation in close relationships. *Emotion Review*, *5*(2), 202–210. https://doi.org/10.1177/1754073912451630

Butner, J., Diamond, L. M., & Hicks, A. M. (2007). Attachment style and two forms of affect coregulation between romantic partners. *Personal Relationships*, *14*(3), 431–455. https://doi.org/10.1111/j.1475-6811.2007.00164.x

Butner, J. E., Crenshaw, A. O., Munion, A. K., Wong, A., & Baucom, B. R. (2018). Simplifying the complexity of interpersonal emotion dynamics with regression graphics. *Studies in emotion and social interaction-Second series. Interpersonal emotion dynamics in close relationships* (pp. 47–74). Cambridge University Press.

Butner, J. E., Deits-Lebehn, C., Crenshaw, A. O., Wiltshire, T. J., Perry, N. S., Kent de Grey, R. G., Hogan, J. N., Smith, T. W., Baucom, K. J., & Baucom, B. R. (2017). A multivariate dynamic systems model for psychotherapy with more than one client. *Journal of Counseling Psychology*, *64*(6), 616–625. https://doi.org/10.1037/cou0000238

Bylsma, L. M., Morris, B. H., & Rottenberg, J. (2008). A meta-analysis of emotional reactivity in major depressive disorder. *Clinical Psychology Review*, *28*(4), 676–691. https://doi.org/10.1016/j.cpr.2007.10.001

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

Castonguay, L. G., & Hill, C. E. (2012). *Transformation in psychotherapy: Corrective experiences across cognitive behavioral, humanistic, and psychodynamic approaches*. American Psychological Association. https://doi.org/10.1037/13747-000

Chaspari, T., Han, S. C., Bone, D., Timmons, A. C., Perrone, L., Margolin, G., & Narayanan, S. S. (2017). Quantifying regulation mechanisms in dating couples through a dynamical systems model of acoustic and physiological

arousal. *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) New Orleans, LA, USA* (pp. 3001–3005).

Cole, P. M., Martin, S. E., & Dennis, T. A. (2004). Emotion regulation as a scientific construct: Methodological challenges and directions for child development research. *Child Development*, *75*(2), 317–333. https://doi.org/10.1111/j.1467-8624.2004.00673.x

Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. F. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech Communication*, *71*, 10–49. https://doi.org/10.1016/j.specom.2015.03.004

Dixon-Gordon, K. L., Bernecker, S. L., & Christensen, K. (2015). Recent innovations in the field of interpersonal emotion regulation. *Current Opinion in Psychology*, *3*, 36–42. https://doi.org/10.1016/j.copsyc.2015.02.001

Duan, C., & Kivlighan, D. M. (2002). Relationships among therapist presession mood, therapist empathy, and session evaluation. *Psychotherapy Research*, *12*(1), 23–37. https://doi.org/10.1080/713869615

Feldman, R. (2012). Bio-Behavioral synchrony: A model for integrating biological and microsocial behavioral processes in the study of parenting. *Parenting: Science and Practice*, *12*(2–3), 154–164. https://doi.org/10.1080/15295192.2012.683342

Feldman, R. (2015). Mutual influences between child emotion regulation and parent–child reciprocity support development across the first 10 years of life: Implications for developmental psychopathology. *Development and Psychopathology*, *27*(4pt1), 1007–1023. https://doi.org/10.1017/S0954579415000656

Fisher, H., Atzil-Slonim, D., Bar-Kalifa, E., Rafaeli, E., & Peri, T. (2016). Emotional experience and alliance contribute to therapeutic change in psychodynamic therapy. *Psychotherapy*, *53*(1), 105–116. https://doi.org/10.1037/pst0000041

Fisher, H., Atzil-Slonim, D., Bar-Kalifa, E., Rafaeli, E., & Peri, T. (2019). Growth curves of clients' emotional experience and their association with emotion regulation and symptoms. *Psychotherapy Research*, *29*(4), 463–478. https://doi.org/10.1080/10503307.2017.1411627

Fonagy, P., Gergely, G., & Jurist, E. L. (Eds.). (2018). *Affect regulation, mentalization and the development of the self*. Routledge. https://doi.org/10.4324/9780429471643

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

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

Gaume, J., Hallgren, K. A., Clair, C., Schmid Mast, M., Carrard, V., & Atkins, D. C. (2019). Modeling empathy as synchrony in clinician and patient vocally encoded emotional arousal: A failure to replicate. *Journal of Counseling Psychology*, *66*(3), 341–350. https://doi.org/10.1037/cou0000322

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment*, *26*(1), 41–54. https://doi.org/10.1023/B:JOBA.0000007455.08539.94

Greenberg, L. S. (2012). Emotions, the great captains of our lives: Their role in the process of change in psychotherapy. *American Psychologist*, *67*(8), 697–707. https://doi.org/10.1037/a0029858

Gross, J. J. (1999). Emotion regulation: Past, present, future. *Cognition and Emotion*, *13*(5), 551–573. https://doi.org/10.1080/026999399379186

Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry*, *26*(1), 1–26. https://doi.org/10.1080/1047840X.2014.940781

Ham, J., & Tronick, E. (2009). Relational psychophysiology: Lessons from mother–infant physiology research on dyadically expanded states of consciousness. *Psychotherapy Research*, *19*(6), 619–632. https://doi.org/10.1080/10503300802609672

Helm, J. L., Sbarra, D., & Ferrer, E. (2012). Assessing cross-partner associations in physiological responses via coupled oscillator models. *Emotion*, *12*(4), 748–762. https://doi.org/10.1037/a0025036

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

Joormann, J., & Stanton, C. H. (2016). Examining emotion regulation in depression: A review and future directions. *Behaviour Research and Therapy*, *86*, 35–49. https://doi.org/10.1016/j.brat.2016.07.007

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*(5), 770–814. https://doi.org/10.1037/0033-2909.129.5.770

Juslin, P. N., & Scherer, K. R. (2005). Vocal expression of affect. In J. A. Harrigan, R. Rosenthal, & K. R. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research* (pp. 65–135). Series in Affective Science Oxford University Press.

Kleinbub, J. R. (2017). State of the art of interpersonal physiology in psychotherapy: A systematic review. *Frontiers in Psychology*, *8*, 2053. https://doi.org/10.3389/fpsyg.2017.02053

Kohut, H. (1971). *The analysis of the self: A systematic approach to the psychoanalytic treatment of narcissistic personality disorders*. University of Chicago Press.

Koole, S. L., & Tschacher, W. (2016). Synchrony in psychotherapy: A review and an integrative framework for the therapeutic alliance. *Frontiers in Psychology*, *7*, 862. https://doi.org/10.3389/fpsyg.2016.00862

Laufer-Goldshtein, B., Talmon, R., & Gannot, S. (2018a). Source counting and separation based on simplex analysis. *IEEE Transactions on Signal Processing*, *66*(24), 6458–6473. https://doi.org/10.1109/TSP.2018.2876349

Laufer-Goldshtein, B., Talmon, R., & Gannot, S. (2018b). Diarization and separation based on a data-driven simplex. In *2018 26th European Signal Processing Conference (EUSIPCO)* (pp. 842–846). IEEE.

Laufer-Goldshtein, B., Talmon, R., & Gannot, S. (submitted). Audio source separation by activity probability detection with maximum correlation and simplex geometry. *EURASIP Journal on Audio Speech and Music*.

Lee, C. C., Black, M., Katsamanis, A., Lammert, A. C., Baucom, B. R., Christensen, A., Georgiou, P. G., & Narayanan, S. S. (2010). Quantification of prosodic entrainment in affective spontaneous spoken interactions of married couples. Eleventh Annual Conference of the International Speech Communication Association, Japan, September 26–30, 2010.

Levitan, R., & Hirschberg, J. B. (2011). Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions. *Proceedings of INTERSPEECH, Florence, Italy, August 27–31, 2011* (pp. 3081–3084).

Lubold, N., & Pon-Barry, H. (2014). Acoustic-Prosodic entrainment and rapport in collaborative learning dialogues. *Proceedings of the 2014 ACM workshop on Multimodal Learning Analytics Workshop and Grand Challenge, Istanbul, Turkey, November, 2014* (pp. 5–12). ACM.

Marci, C. D., Ham, J., Moran, E., & Orr, S. P. (2007). Physiologic correlates of perceived therapist empathy and social-emotional process during psychotherapy. *Journal of Nervous and Mental Disease*, *195*(2), 103–111. https://doi.org/10.1097/01.nmd.0000253731.71025.fc

McCullough, L. (Ed.). (2003). *Treating affect phobia: A manual for short-term dynamic psychotherapy*. Guilford Press.

Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., & Claud, D. A. (2003). The outcome rating scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy*, *2*(2), 91–100.

Mitchell, S. A. (1993). *Hope and dread in psychoanalysis*. Basic Books.

Mundt, J. C., Vogel, A. P., Feltner, D. E., & Lenderking, W. R. (2012). Vocal acoustic biomarkers of depression severity and treatment response. *Biological Psychiatry*, *72*(7), 580–587. https://doi.org/10.1016/j.biopsych.2012.03.015

Pos, A. E., Paolone, D. A., Smith, C. E., & Warwar, S. H. (2017). How does client expressed emotional arousal relate to outcome in experiential therapy for depression? *Person-Centered and Experiential Psychotherapies*, *16*(2), 173–190. https://doi.org/10.1080/14779757.2017.1323666

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, *31*(4), 437–448. https://doi.org/10.3102/10769986031004437

Radkovsky, A., McArdle, J. J., Bockting, C. L., & Berking, M. (2014). Successful emotion regulation skills application predicts subsequent reduction of symptom severity during treatment of major depressive disorder. *Journal of Consulting and Clinical Psychology*, *82*(2), 248–262. https://doi.org/10.1037/a0035828

Rafaeli, E., Bernstein, D. P., & Young, J. (2010). *Schema therapy: Distinctive features*. Routledge. https://doi.org/10.4324/9780203841709

Ramseyer, F., & Tschacher, W. (2014). Nonverbal synchrony of head-and body-movement in psychotherapy: Different signals have different associations with outcome. *Frontiers in Psychology*, *5*, 979. https://doi.org/10.3389/fpsyg.2014.00979

Ramseyer, F. T. (2020). Exploring the evolution of nonverbal synchrony in psychotherapy: The idiographic perspective provides a different picture. *Psychotherapy Research*, *30*(5), 622–634.

Reed, R. G., Barnard, K., & Butler, E. A. (2015). Distinguishing emotional coregulation from codysregulation: An investigation of emotional dynamics and body weight in romantic couples. *Emotion*, *15*(1), 45–60. https://doi.org/10.1037/a0038561

Rice, L., & Wagstaff, A. (1967). Client voice quality and expressive style as indexes of productive psychotherapy. *Journal of Consulting Psychology*, *31*, 557–563. https://doi.org/10.1037/h0025164

Rochman, D., & Amir, O. (2013). Examining in-session expressions of emotions with speech/vocal acoustic measures: An introductory guide. *Psychotherapy Research*, *23*(4), 381–393. https://doi.org/10.1080/10503307.2013.784421

Sbarra, D. A., & Hazan, C. (2008). Coregulation, dysregulation, self-regulation: An integrative analysis and empirical agenda for understanding adult attachment, separation, loss, and recovery. *Personality and Social Psychology Review*, *12*(2), 141–167. https://doi.org/10.1177/1088868308315702

Schuller, B., Batliner, A., Steidl, S., & Seppi, D. (2011). Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge. *Speech Communication*, *53*(9–10), 1062–1087. https://doi.org/10.1016/j.specom.2011.01.011

Shapira, N., Lazarus, G., Goldberg, Y., Gilboa-Schechtman, E., Tuval-Mashiach, R., Juravski, D., & Atzil-Slonim, D. (2020). Using computerized text analysis to examine associations between linguistic features and clients' distress during psychotherapy. *Journal of Counseling Psychology*. Advance online publication. https://doi.org/10.1037/cou0000440

Shedler, J. (2010). The efficacy of psychodynamic psychotherapy. *American Psychologist*, *65*(2), 98–109. https://doi.org/10.1037/a0018378

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (MINI): 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.

Sheppes, G., Suri, G., & Gross, J. J. (2015). Emotion regulation and psychopathology. *Annual Review of Clinical Psychology*, *11*, 379–405. https://doi.org/10.1146/annurev-clinpsy-032814-112739

Sloan, D. M., & Kring, A. M. (2007). Measuring changes in emotion during psychotherapy: Conceptual and methodological issues. *Clinical Psychology: Science and Practice*, *14*(4), 307–322. https://doi.org/10.1111/j.1468-2850.2007.00092.x

Soma, C. S., Baucom, B. R., Xiao, B., Butner, J. E., Hilpert, P., Narayanan, S., Atkins, D. C., & Imel, Z. E. (2019). Coregulation of therapist and client emotion during psychotherapy. *Psychotherapy Research*, 1–13.

Summers, R. F., & Barber, J. P. (2009). *Psychodynamic therapy: A guide to evidence-based practice*. Guilford Press.

Tellegen, A., Watson, D., & Clark, L. A. (1999). On the dimensional and hierarchical structure of affect. *Psychological Science*, *10*(4), 297–303. https://doi.org/10.1111/1467-9280.00157

Thompson, R. A. (2011). Emotion and emotion regulation: Two sides of the developing coin. *Emotion Review*, *3*(1), 53–61. https://doi.org/10.1177/1754073910380969

Tomicic, A., Martinez, C., & Krause, M. (2015). The sound of change: A study of the psychotherapeutic process embodied in vocal expression. Laura Rice's ideas revisited. *Psychotherapy Research*, *25*(2), 263–276. https://doi.org/10.1080/10503307.2014.892647

Tschacher, W., & Meier, D. (2020). Physiological synchrony in psychotherapy sessions. *Psychotherapy Research*, *30*(5), 558–573.

Tschacher, W., Rees, G. M., & Ramseyer, F. (2014). Nonverbal synchrony and affect in dyadic interactions. *Frontiers in Psychology*, *5*, 1323. https://doi.org/10.3389/fpsyg.2014.01323

Uchino, B. N., Cacioppo, J. T., & Kiecolt-Glaser, J. K. (1996). The relationship between social support and physiological processes: A review with emphasis on underlying mechanisms and implications for health. *Psychological Bulletin*, *119*(3), 488–531. https://doi.org/10.1037/0033-2909.119.3.488

Valstar, M., Gratch, J., Schuller, B., Ringeval, F., Lalanne, D., Torres, M., Scherer, S., Stratou, G., Cowie, R., & Pantic, M. (2016). Avec 2016: Depression, mood, and emotion recognition workshop and challenge. *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge United States* (pp. 3–10). ACM.

Winnicott, D. W. (1971). *Playing and reality*. Penguin.

Yang, Y., Fairbairn, C., & Cohn, J. (2013). Detecting depression severity from vocal prosody. *IEEE Transactions on Affective Computing*, *4*, 142–150. https://doi.org/10.1109/T-AFFC.2012.38

Zaki, J., & Williams, W. C. (2013). Interpersonal emotion regulation. *Emotion*, *13*(5), 803–810. https://doi.org/10.1037/a0033839

Received March 26, 2020
Revision received October 14, 2020
Accepted December 14, 2020 ■