Empirical Article

Clinical Psychological Science
2024, Vol. 12(5) 840–854

DOI: 10.1177/21677026231195793
www.psychologicalscience.org/CPS

# Facing Emotions: Between- and Within-Sessions Changes in Facial Expression During Psychological Treatment for Depression

**Hadar Fisher[1], Philip T. Reiss[2], Dovrat Atias[1], Michal Malka[1], Ben Shahar[3], Simone Shamay-Tsoory[1], and Sigal Zilcha-Mano[1]**

[1]Department of Psychology, University of Haifa; [2]Department of Statistics, University of Haifa; and [3]School of Social Work and Social Welfare, The Hebrew University of Jerusalem

## Abstract

The main diagnostic criteria for major depressive disorder (MDD) are consistent experiences of high levels of negative emotions and low levels of positive emotions. Therefore, modification of these emotions is essential in the treatment of MDD. In the current study, we harnessed a computational approach to explore whether experiencing negative emotions during psychological treatment is related to subsequent changes in these emotions. Facial expressions were automatically extracted from 175 sessions of 58 patients with MDD. Within sessions, a U-shaped trajectory of change in valence was observed in which patients expressed an increase in negative emotions in the middle of the session. Between sessions, a consistent increase in valence was observed. A trajectory of within-sessions decrease followed by an increase in valence was positively associated with greater perceived positive emotions and subsequent decreases in depressive symptoms. These findings highlight the importance of targeting negative emotions during treatment to achieve more favorable outcomes.

## Keywords

major depressive disorder, emotional avoidance, negative emotions, positive emotions, facial expression

Received 9/4/22; Revision accepted 7/24/23

Depression is one of the most prevalent mental-health disorders and the leading cause of disability worldwide (American Psychiatric Association, 2013; Friedrich, 2017). Much of the impairment in the functioning of individuals diagnosed with MDD emanates from mood disturbances in both positive and negative emotions (Gross & Jazaieri, 2014). Although there is an ongoing debate in the field as to whether individuals with MDD consistently display both high levels of negative emotions and low levels of positive emotions, these symptoms are currently used to diagnose depression (Hamilton, 1967; Joormann & Stanton, 2016; Joormann & Vanderlind, 2014). Although a range of empirically supported interventions are available for the treatment of MDD, rates of nonresponse remain high, highlighting the pressing need to improve current treatments (Cuijpers et al., 2014). One way to improve treatment is by targeting the fundamental processes that go awry in depression.

One of the leading theoretical models to understand the mechanisms underlying MDD suggests that the emotional impairment observed in MDD results from "emotional avoidance." Emotional avoidance is defined as an unhealthy attempt to control negative emotions by avoiding them (S. C. Hayes et al., 1996). According to this model, in an attempt to regulate distress, individuals with MDD exert considerable efforts to suppress and avoid the experience of negative emotions. This

**Corresponding Author:**
Hadar Fisher, Department of Psychology, University of Haifa
Email: hadar.fisher@gmail.com

suppression prevents them from effectively responding to emotional stimuli and thus may reduce emotional distress in the short term but increases emotional distress and attenuates positive emotionality over time (Gross, 2013). Hence, emotional avoidance may increase the risk for depression (Bylsma et al., 2008). Consistent with the conceptualization of emotional avoidance as a mechanism underlying symptoms of MDD, numerous studies have found emotional avoidance to be related to depressive symptoms (Campbell-Sills et al., 2006; Cribb et al., 2006; Forbes et al., 2020; Mellick et al., 2019; Tull et al., 2004). In addition, strategies of expressing emotions were found to be associated with a decrease in depressive symptoms (Kahn & Garrison, 2009).

To improve mental health, most psychological treatment modalities incorporate interventions aimed at exposing patients to previously suppressed emotions and helping them experience and express them in adaptive ways (for reviews, see Aafjes-van Doorn & Barber, 2017; Greenberg, 2012; Summers & Barber, 2010; Thoma & McKay, 2014). It is assumed that people often avoid their emotions because they are unable to regulate them. Through therapy, people can first express their emotions and then learn how to regulate and process them in a healthier manner (Greenberg, 2012). Thus, negative emotions are expected to increase during a session but decrease from session to session as the patients enhance their ability to handle these emotions. Previous studies have found that interventions aimed at decreasing emotional avoidance and encouraging patients' greater experience of emotions were associated with symptom decrease and improved mental functioning (Diener et al., 2007; Fisher et al., 2020; Peluso & Freund, 2018). In addition, patients who experienced their emotions during treatment and did it in a stable manner showed an improvement in their ability to regulate their emotions (Fisher et al., 2019). However, scant attention has been paid to the process of emotion modification—in which negative emotions are transformed into positive ones—during psychological treatment. In the current study, we thus explored how emotions change both within and between sessions by measuring emotions as they fluctuate continuously within each session. We also explored whether targeting emotional avoidance through the experience of negative emotions during psychological treatment can change the way emotions are experienced by patients diagnosed with MDD and improve depressive symptoms.

Most previous studies have identified emotions by self-report questionnaires or observer ratings, both of which curtail researchers' ability to repeatedly measure within-sessions change (Sloan & Kring, 2007). Self-reports provide only a retrospective snapshot and thus cannot fully capture emotional dynamics during a session. Observer ratings are time-consuming, which makes them impractical for data with repeated observations for each participant. In addition, previous studies in psychotherapy have focused on patients' verbal responses or self-reports to detect the experience of emotions; however, emotions are often also expressed nonverbally, through tone of voice or facial expressions. Therefore, focusing on the nonverbal expression of emotions can provide a deeper understanding of the ways in which emotions are displayed and modified during psychological treatment. Facial expressions constitute one of the most fundamental types of nonverbal expressions of emotions.

Depression as a mood disorder can be portrayed through a variety of nonverbal signs, including facial expression (Ellgring, 1989; Waxer, 1974). For example, the duration and intensity of spontaneous smiles (Scherer et al., 2013, 2014) or a lack of smiles (Lucas et al., 2015) are considered to contain valuable patterns for depression detection. Previous studies using automatic coding of facial expressions have reported that variance in facial expressions successfully predicted the presence of depression (Alghowinem et al., 2015; Girard et al., 2014; Pampouchidou et al., 2016; Schneider et al., 1990; Stratou et al., 2015). However, very little is known about how facial expressions change during the treatment of depression and how this change is related to patients' perceived emotions and to changes in depressive symptoms.

Recent research has drawn on computer vision and machine learning to automatically identify emotions from video recordings of facial expressions (Harley, 2016; Ko, 2018). Recent developments in computer science have significantly improved the validity, reliability, and accessibility of automated facial-expression analysis (Cootes et al., 2001; Lewinski et al., 2014; Sandbach et al., 2012). These technologies are noninvasive, allowing for the detection of spontaneous facial behavior without the use of recording equipment, such as electrodes and wires. A growing body of literature has explored the reliability of automatic facial coding in detecting emotions by comparing it with traditional methods. Results have shown that automatic facial coding generates values that are comparable and highly correlated with those generated by untrained human coders (Krumhuber et al., 2021a, 2021b), trained human coders (Girard & Cohn, 2015; Girard et al., 2013; Gupta et al., 2022), and electromyography, which is currently considered the psychophysiological "gold standard" (Beringer et al., 2019; Höfling et al., 2021; Kulke et al., 2020).

Although there is growing use of these automated methods in psychology (Haines et al., 2019), few works have employed this method in psychotherapy (Arango et al., 2019). To the best of our knowledge, only one study has collected and analyzed facial expressions to explore how emotions change during therapy. In this study, facial expressions were automatically coded at the beginning and end (first 10 min and last 10 min) of 29 therapy sessions of 12 patients. A significant decrease in negative valence was found from the beginning of the sessions to the end (Arango et al., 2019). These findings point to the importance of psychological treatment in modifying emotions. However, given that only the beginnings and ends of the sessions were analyzed, it is not clear how emotions changed throughout the sessions. In most therapy sessions, patients share their difficulties with the therapists, thus suggesting that the increase in positive emotions found by Arango et al. (2019) may have occurred after an increase in negative emotions during the middle of the session, generating a U-shape of decrease followed by an increase.

In the current study, we incorporated automatic facial-recognition tools to explore how nonverbal emotional expressions change over the course of treatment. We explored changes in the short term, within the therapy session, and in the long term, between therapy sessions. We expected that a temporary increase in negative emotions would be observed within the therapy session but that there would be consistent decrease in negative emotions and increase in positive emotions over the course of treatment. To explore the gradual transition from positive to negative emotions and vice versa within a session, we used a dimensional approach according to which emotions exist on a continuum of affective dimensions (Barrett, 1998). In the dimensional approach, transitions of emotions can be easily captured on one or more continuous scales (Gunes & Schuller, 2013). One such scale is the “valence” scale, which refers to whether an emotion is pleasant/positive or unpleasant/negative and rated continuously from negative (low valence) to positive (high valence) emotions. The dimensional approach is consistent with previous research on depression, which showed that it is related to various positive and negative emotions (Power & Tarsia, 2007; Watson et al., 2011). In psychotherapy for depression, one of the main goals is to process, regulate, and transform negative emotions regardless of their specific category (Greenberg & Safran, 1990). This approach is widely used in depression research and has been applied in cases of human coding (e.g., Girard et al., 2013), automatic facial-expression coding (Flynn et al., 2020; Girard et al., 2013; Stöckli et al., 2018), physiological measures (functional MRI; Habes et al., 2013; Kim et al., 2020), and self-reports (Kuppens et al., 2013).

Finally, we explored how change in these expressions converged with patients’ prospective reports on their in-session emotions. On the basis of the literature on emotional avoidance, we hypothesized the following:

*Hypothesis 1:* Within treatment sessions, patients will exhibit a pattern of decrease in valence (more negative and fewer positive emotions) followed by an increase in valence (more positive and fewer negative emotions).

*Hypothesis 2:* Between sessions, patients will show an increase in positive valence and a decrease in negative valence from session to session.

*Hypothesis 3:* Within sessions, a pattern of decrease followed by an increase in valence will be associated with patients reporting that they experienced fewer negative emotions and more positive emotions during the session.

In addition, as an exploratory hypothesis, we investigated whether in-session changes in valence would also predict subsequent decreases in depressive symptoms.

## Transparency and Openness

The study analysis code and supplemental materials are available at https://osf.io/3mehp. In addition, Supplemental Material is available online. When this study was carried out, the informed-consent form for the participants stated that we would keep the data strictly confidential and would not be shared. Therefore, the data are not available. We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. The study design, procedure, and informed-consent form were approved by the Ethics Committee of the University of Haifa (Approval No. 395/19, October 30, 2019).

## Method

### *Patients and procedure*

The initial sample included 86 patients enrolled in 16-session manualized psychotherapy for MDD as part of a randomized control trial (RCT) comparing supportive treatment and supportive-expressive treatment (Zilcha-Mano et al., 2021). The inclusion criteria for participating in the RCT were (a) a current MDD diagnosis based on the Mini International Neuropsychiatric Interview (Sheehan et al., 1998) and scores above 14 on the 17-item Hamilton Rating Scale for Depression (HRSD; Hamilton, 1967) on two evaluations, 1 week apart; (b) if on medication, patients’ dosage had to be

**Table 1.** Patients' Demographic and Clinical Characteristics

| | *N* | Percent | *M* (*SD*) |
|---|---|---|---|
| Age | | | 32.03 (8.54) |
| Gender | | | |
| Female | 51 | 59.3 | |
| Male | 35 | 40.7 | |
| Years of education | | | 14.83 (2.90) |
| Ethnicity/religion | | | |
| Jewish | 69 | 80.23 | |
| Christian | 9 | 10.46 | |
| Muslim | 5 | 5.81 | |
| Atheists | 3 | 3.49 | |
| Marital status | | | |
| Single | 64 | 74.4 | |
| Married | 15 | 17.4 | |
| In a relationship | 2 | 2.3 | |
| Divorced/widow | 5 | 5.9 | |
| Comorbidity | | | |
| Any personality disorder | | 70.9 | |
| Obsessive compulsive | | 46.5 | |
| Avoidant | | 30.2 | |
| Dependent | | 16.3 | |
| Borderline | | 15.1 | |
| Narcissistic | | 12.8 | |
| Histrionic | | 5.8 | |
| Any anxiety disorder | | 72.1 | |
| Generalized anxiety disorder | | 60.5 | |
| Social anxiety disorder | | 30.2 | |
| Panic disorder | | 16.3 | |
| Agoraphobia | | 16.3 | |
| Obsessive compulsive disorder | | 4.7 | |
| Posttraumatic stress disorder | | 10.5 | |

stable and they had to be willing to maintain a stable dosage for the duration of treatment; (c) age from 18 to 60 years; and (d) Hebrew-language fluency. The exclusion criteria were (a) current risk of suicide or self-harm (HRSD suicide item > 2); (b) current substance abuse disorder; (c) current or past schizophrenia or psychosis, bipolar disorder, or severe eating disorder, requiring medical monitoring; (d) history of organic mental disease; and (e) currently in psychotherapy. Table 1 lists the patients' demographic and clinical characteristics. Further information on the RCT is provided in the Supplemental Material.

During the active phase of treatment, patients completed the Positive and Negative Affect Scale (PANAS; Watson et al., 1988) immediately after each session. All sessions were videotaped in the consulting room. The participants were given comprehensive information about the study procedure both orally and in writing. This information covered the possible implications of their participation, including potential risks, inconveniences, and benefits. Patients willing to participate signed informed-consent forms confirming their understanding that all treatment sessions would be videotaped and that they would have the right to withdraw from the study at any time.

The therapy sessions were recorded using a Type IP PTZ HD 1080-pixel (px) camera with a resolution of 1080 px (1920 × 1080 px) at a standard frame rate of 25 frames per second. The camera was attached to the wall in front of the chair in two therapy rooms where the chairs were fixed in place. The camera angle relative to the patients' faces was frontal and at eye level such that the camera was not able to move or zoom during the sessions.

## Measures

***Automatic coding of facial expression.*** The video recordings from Sessions 1, 4, 8, 12, and 16 were processed by automated analysis using the AFFDEX classifier, which is part of the iMotions Biometric Research Platform (Version 9.3; iMotions, 2022). The Affectiva

AFFDEX system is a widely used software-development kit that classifies emotional states (Magdin et al., 2019). Similar to other automated facial-expression-analysis systems, AFFDEX uses the facial action-coding system (FACS; Ekman & Rosenberg, 2005), an objective taxonomy widely used for coding facial behavior. FACS comprises approximately 33 anatomically based facial actions, known as action units (AUs), which combine to create various facial expressions. The system can recognize basic emotions through a combination of 20 AUs, which interact to form different facial expressions. It estimates the probability of occurrence (in percentages) for each of the seven basic emotions (anger, sadness, contempt, disgust, fear, joy, and surprise) within 40-ms time windows, providing scores of 0 to 100 for each frame (McDuff et al., 2016). More information about this system is provided in the Supplemental Material. For the purposes of the current study, we used the valence as analyzed by iMotions software. The valence metric, also calculated from a set of observed facial expressions, indicates whether the emotional state of the participant is positive or negative with a range of values from –100 to 100. Recent research suggests that valence is preferable when using automatic coding of facial expressions because it can resolve the issue of lower accuracy rates for nonprototypical emotions (Yitzhak et al., 2017). Valence can be applied to all relevant expressions, including subtle and complex emotions that may be missed with a categorical approach (Stöckli et al., 2018). In addition, the range of –100 to 100 is less subject to floor and ceiling effects than emotion variables that range from 0 to 100 and is closer to the normal distribution conventionally assumed in linear mixed-effect models. Information about the validity of the AFFDEX valence estimates in our data set is provided in the Supplemental Material.

***Validity of the AFFDEX valence estimates.*** A total of 24 sessions (2,132 observations) were coded by four raters who had no prior training. Eight sessions were coded by one rater, 15 were coded by two raters, and one was coded by three raters. The advantage of using untrained volunteers is that it allows for a more naturalistic approach to studying how people perceive facial expressions rather than relying on experts who may have preconceptions or biases (Tottenham et al., 2009). In our analyses, AFFDEX ratings were averaged for every 30 s. To match this, the raters observed video recordings of the sessions and evaluated the valence of each 30-s interval on a scale ranging from –100 to 100. Thus, half-minutes were nested within sessions. There is no additional level of session within subjects because we evaluated no more than one session for each subject.

Validity of the AFFDEX estimates for our data was assessed in two stages. First, we calculated the interrater reliability among the human raters. Second, we examined the consistency between the human ratings (averaged across raters) and the ratings obtained from AFFDEX. Interrater reliability for the human raters was measured by intraclass correlation (ICC) for multilevel data using maximum likelihood estimation (Ten Hove et al., 2022). According to standard benchmarks for κ or ICC (Fleiss et al., 2013), which consider values of .40 and below as poor, .40 to .59 as fair, .60 to .74 as good, and .75 or above as excellent, we found fair interrater reliability at the within-sessions level (ICC = .52) and excellent interrater reliability at the patient level (ICC = .81).

To explore the consistency between human and automatic coding, we conducted correlation analyses at both the within-sessions and between-patients levels. Within-sessions correlation was computed by the method of Bland and Altman (1995) implemented in R by Bakdash and Marusich (2017). At the between-patients level, we estimated the overall correlations between average session ratings. The correlation between the valence derived from human coding and automated coding was strong at the within-sessions level ($r = .38$, $p < .001$) and stronger at the between-patients level ($r = .68$, $p < .001$). These findings align with a recent study by Gupta et al. (2022), which also demonstrated a substantial association between iMotions coding and human coding.

***The PANAS.*** The PANAS (Watson et al., 1988) was used to assess patients' retrospective reports on their emotions during each therapy session. The PANAS is a self-report questionnaire consisting of 20 items, 10 assessing subjective experience of positive affect (PA; e.g., interested, enthusiastic) and 10 assessing subjective experience of negative affect (NA; e.g., irritable, upset). Patients were asked to rate the extent to which they experienced each of the presented emotions during the session. Items were rated on a scale from 1 (*very low or not at all*) to 5 (*very much*). A mean score of positive and negative emotions was calculated for each session throughout treatment. Internal consistency is usually considered acceptable if the estimate is .70 or higher (A. F. Hayes & Coutts, 2020; McNeish, 2018; for a discussion of beliefs about conventional cutoffs for "acceptable" reliability, see Lance et al., 2006). Reliability in our study was acceptable, with a Cronbach's α of .92 and a McDonald's ω of .94 for PA and Cronbach's α of .88 and a McDonald's ω of .92 for NA.

***The HRSD.*** To evaluate symptomatic change, the HRSD (Hamilton, 1960), a semistructured clinical interview designed to assess the severity of depression, was administered at baseline and before each session. The HRSD contains 21 items rated on 3-point (0–2) or 5-point (0–4) scales. The total score consists of the sum of the responses and ranges from 0 to 52 points, corresponding to greater severity of depression. The internal reliability was found to be acceptable, with a Cronbach's $\alpha$ of .83 and a McDonald's $\omega$ of .85.

## Data preprocessing

To ensure the ecological validity and to avoid altering the usual routine of the session or influencing the interaction between the patient and the therapist, patients were not instructed to look at the camera. As a result, some videos lacked sufficient quality to be processed given the loss of data continuity because of the lack of follow-up of the patient's face. To restrict our analyses to relatively high-quality data, we took the following steps.

First, the recordings were divided into segments of 30 s (750 frames), and only segments with at least 10% coded frames (75 frames) were considered "valid" for further analysis. Second, sessions were retained only if they included at least 40 valid 30-s segments, containing a total of at least 15,000 coded frames (i.e., approximately 20% of the session).

This process resulted in 175 naturalistic videos from the therapy sessions of 58 patients (14 patients with all five sessions, 11 with four sessions, 10 with three sessions, nine with two sessions, and 13 with one session). A comparison of these 175 retained sessions to the other 172 sessions (those with insufficient available coded frames) in terms of postsession PANAS (Watson et al., 1988) and presession depressive symptoms using the HRSD (Hamilton, 1960) indicated no significant differences (PANAS negative scale: $t = -0.80$, $p = .42$; PANAS positive scale: $t = 0.05$, $p = .96$; HRSD: $t = -0.48$, $p = .63$). To check the sensitivity of the results to different thresholds, we reran the models described below with three alternatives for each threshold. These alternative threshold values did not significantly affect the model results.

## Data analytic plan

Therapy sessions were not equal in their length, although most of the sessions followed the 50-min protocol ($M = 53.2$ min, $SD = 3.9$, range = 44–60). To standardize the session length, we divided the time in session by the length of the session. Thus, the new variable range was between 0 and 1. For reasons related to interpretation of the models (see explanation after Model 2 below), we subtracted 1 from the session numbers (i.e., the sessions analyzed were numbered 0, 3, 7, 11, and 15).

Because of the scant literature on within-sessions changes in emotions during treatment, we adopted a flexible data-driven approach to describe within- and between-sessions changes in valence rather than assuming a prespecified form, such as linear or quadratic. Specifically, we used spline-based models in the general framework of semiparametric regression (Ruppert et al., 2003; Wood, 2017). This approach allows for flexible shapes of change in variables over time.

Given that assessments were conducted for five sessions ($s = 0, 3, 7, 11, 15$), the most general model that can be estimated for the mean valence at time $t$ within session $s$ is

$$E(v|t,s) = f_s(t), \tag{1}$$

where the mean valence is given by distinct smooth curves $f_0(t), f_3(t), f_7(t), f_{11}(t), f_{15}(t)$ for sessions 0, 3, 7, 11, and 15, respectively.

More restrictive alternatives to Model 1 may be formulated by assuming (a) that the mean between-sessions change be a linear function of session number $s$ for each within-sessions time point $t$ and/or (b) that the mean between-sessions change be a constant independent of $t$, implying that the shape of the within-sessions mean valence trajectory is the same for each session.

The first requirement leads to the varying-coefficient model (Hastie & Tibshirani, 1993)

$$E(v|t,s) = f_0(t) + \beta(t)s, \tag{2}$$

where $\beta(t)$ is a smooth function. In this model, between-sessions change is linear, but the intercept and slope of the line vary flexibly among time points within the session so that the shape of the mean within-sessions valence trajectory can differ across sessions. Note that because the first session is numbered $s = 0$, Model 2 implies that $f_0(t)$ is the mean valence for that session, as in Model 1 and the two additional models below.

The second requirement yields the additive model

$$E(v|t,s) = f_0(t) + \kappa_s, \tag{3}$$

where $\kappa_0 = 0$ and $\kappa_3, \kappa_7, \kappa_{11}, \kappa_{15}$ are constants. This model allows nonlinear between-sessions change, but the mean trajectory has the same shape for each session.

Imposing both requirements leads to the model

$$E(v|t,s) = f_0(t) + \beta s. \tag{4}$$

Note that Model 4 is equivalent to Model 2 with $\beta(t)$ required to be constant or to Model 3 with session treated as continuous rather than categorical.

Models 1 through 4 were fitted using the *mgcv* (Wood, 2017) and *gamm4* (Wood & Scheipl, 2020) packages for R (R Core Team, 2022). These packages apply roughness penalties as a form of regularization. The tuning parameters that determine how much regularization to apply are chosen optimally by likelihood-based methods that are described in Wood (2011) and Wood et al. (2016). This approach allows us to avoid oversmoothing (underfitting) or undersmoothing (overfitting) of the data, ensuring that the models accurately captured the underlying trends in the data. For details on the spline implementation, see the Supplemental Material.

The models included random effects of subject and session within subjects. For example, in Model 4, the valence for subject $i$ at time point $t_{ijk}$ within his or her $j$th observed session (session $s_{ij}$) is

$$v_{ijk} = f_0\left(t_{ijk}\right) + \beta s_{ij} + \zeta_i + \eta_{ij} + \varepsilon_{ijk}, \tag{5}$$

where $\zeta_i, \eta_{ij}$ denote random effects of subject and of session within subject, respectively, and $\varepsilon_{ijk}$ denotes the residual; the random subject effects, random session effects, and residuals are taken as independent normal with mean 0 and variances $\sigma_\zeta^2, \sigma_\eta^2, \sigma_\varepsilon^2$, respectively. The four models were compared with the Akaike information criterion (AIC) as modified by Wood et al. (2016) for smooth models.

AIC was also used to evaluate the necessity of flexible models versus simpler quadratic parametric fits. The results showed that the flexible nonparametric approach fit the data better (see Table S2 in the Supplemental Material).

We fitted an additional model with a more complex random-effects structure to examine the relationship between within- and between-sessions change in valence (for this post hoc analysis, see Results section).

To investigate whether patients' in-session trajectory of valence, as continuously evaluated by their facial expressions, could predict their reports of in-session negative and positive emotions and changes in depressive symptoms, we used scalar-on-function regression models (Dziak et al., 2019; Reiss et al., 2017). These models allow for a data-driven examination of which aspects of the within-sessions dynamics predict the overall assessment of the session by the patient. In their simplest version (omitting random effects), these models are of the form $y = \alpha + \int v(t)\beta(t)dt + \varepsilon$, where $y$ is the outcome (positive PANAS, negative PANAS, or change in HRSD), $v(t)$ denotes valence at time $t$, $\alpha$ is an intercept, $\varepsilon$ denotes error, and $\beta(t)$ is a coefficient function representing the effect of valence at time $t$. This is similar in meaning to a conventional multiple regression model in which the integral would be replaced by a sum $\sum v_t\beta_t$, over discrete time points $t$, with $v_t$ denoting valence at the given time point and $\beta_t$ denoting the corresponding effect. However, we used penalized splines to estimate the model, resulting in a coefficient function $\beta(t)$ that is smooth and thus may be easier to interpret than a discrete collection of coefficients would be. For this type of multilevel data set, the model can be formulated as

$$y_{ij} = \alpha + \int v_{ij}(t)\beta(t)\,dt + u_i + \varepsilon_{ij}, \tag{6}$$

where subscripts $ij$ denote measurements for the $i$th client's $j$th session and $u_i$ denotes the random effect of client $i$. This model was fitted using the R package *refund* (Goldsmith et al., 2022). The complete R code and a supplementary document describing additional analyses are available on OSF at https://osf.io/3mehp.

## Results

When describing the results, a decrease in valence means more negative and less positive valence, whereas an increase in valence means more positive and less negative emotions. As described in more detail above, Model 1 is the most flexible because it allows the within-sessions valence trajectory to differ between sessions and does not assume linear change in the means between sessions. Model 2 also allows the within-sessions valence trajectory to differ between sessions but assumes linear change between the sessions. Model 3 assumes that the valence trajectory is the same for each session but does not assume linear change between the sessions. Model 4 is the least flexible because it assumes the same valence trajectory for each session and linear change between the sessions. In all models, the within-sessions curve $f_0(t)$ was found to differ significantly from zero ($p$ < .001), by the test of Wood (2013). Estimates of the curves $f_0(t), f_3(t), f_7(t), f_{11}(t), f_{15}(t)$ in Model 1 are displayed in Figure 1. These estimates and those in Figures 2 and 3 are accompanied by confidence intervals defined by the estimate ±2 *SE*. As shown in Figure 1, there were some differences in the estimated mean valence curves for the five sessions, but the general pattern was U-shaped such that the average valence decreased in the middle of the session and then rose toward the end. The AIC values for Models 1 through 4 were 131,554.4, 131,550.9, 131,535.5, and 131,541.9, respectively. The fact that Models 1 and 2 had the highest AIC values implies that the data did not support differently shaped within-sessions mean valence curves for different sessions. Consistently, the estimate of $\beta(t)$ in Model 2 was quite flat and very close to the constant estimate 0.256 obtained from Model 4. The estimate of $f_0(t)$ in Model 2 (not shown) is very similar to that of Model 4. For Model 3, the estimate of $f_0(t)$, displayed in Figure 2a, reveals a similar U-shape.

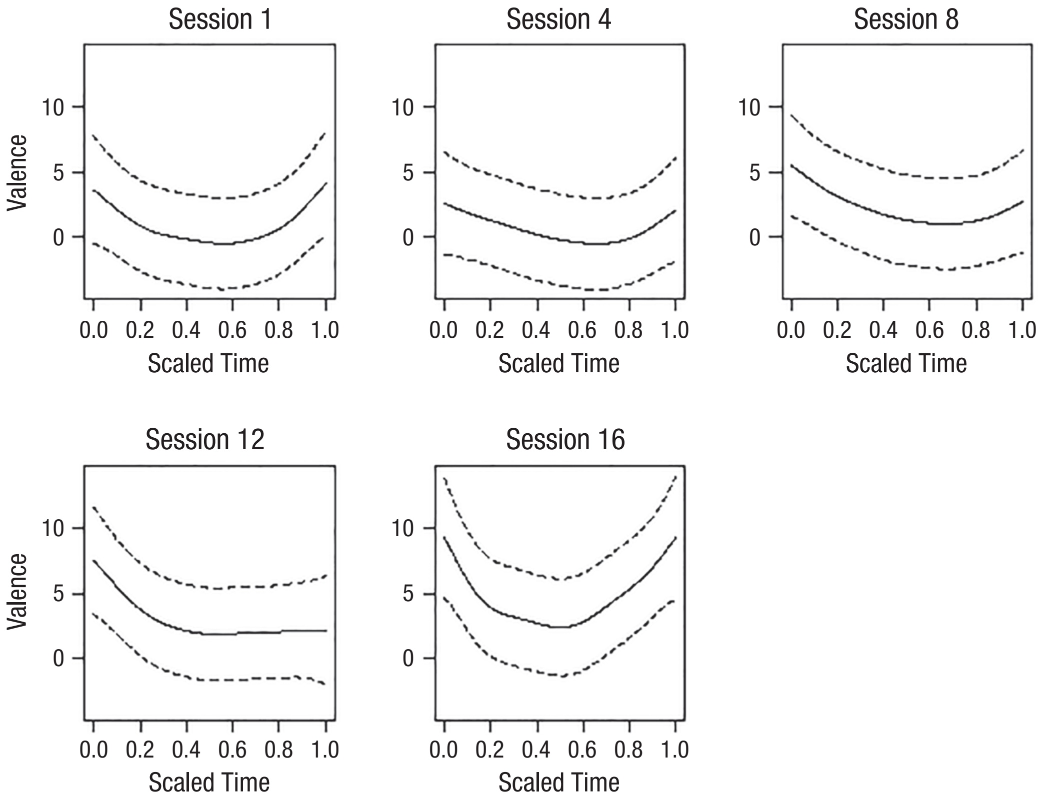


**Fig. 1.** Model 1: curves representing the within-sessions changes in valence for each session. The solid line represents the mean change over the session. The dotted lines represent the confidence interval (CI).

Estimates of the parameters $\kappa_3, \kappa_7, \kappa_{11}, \kappa_{15}$ in this model, representing the constant mean change in valence for follow-up Sessions 3, 7, 11, and 15 relative to Session 0, are presented in Table 2 and in Figure 2b; Table 2 also presents the estimate of Model 4 parameter $\beta$, representing the mean change in valence per session. Like Model 4, Model 3 was consistent with a pattern of between-sessions increase but suggests that it began around Session 3. Additional post hoc analyses are reported in the Supplemental Material.[1]

The above results indicate that on average, the valence followed a U-shaped trajectory within sessions and tended to increase between sessions. This prompted us to investigate whether there was an association between U-shaped within-sessions change and between-sessions increase. To address this question,

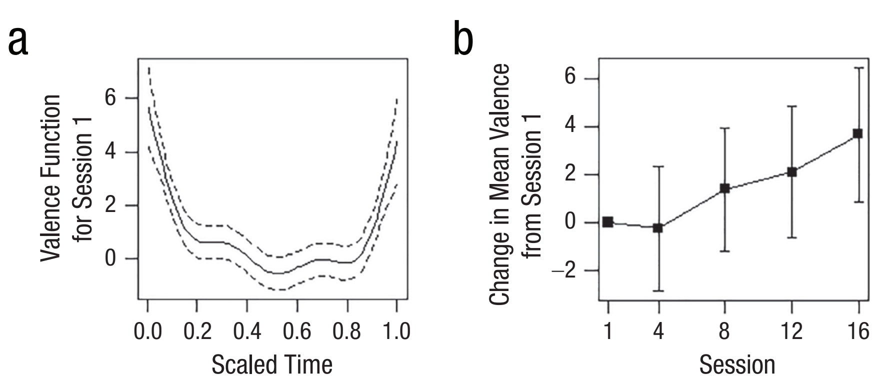


**Fig. 2.** (a) Estimated average valence trajectory within Session 0 according to Model 3. The average valence for each of the later sessions had the same shape according to this model but was shifted vertically by a constant. (b) Estimated average between-sessions change in valence according to Model 3 (see Table 1).

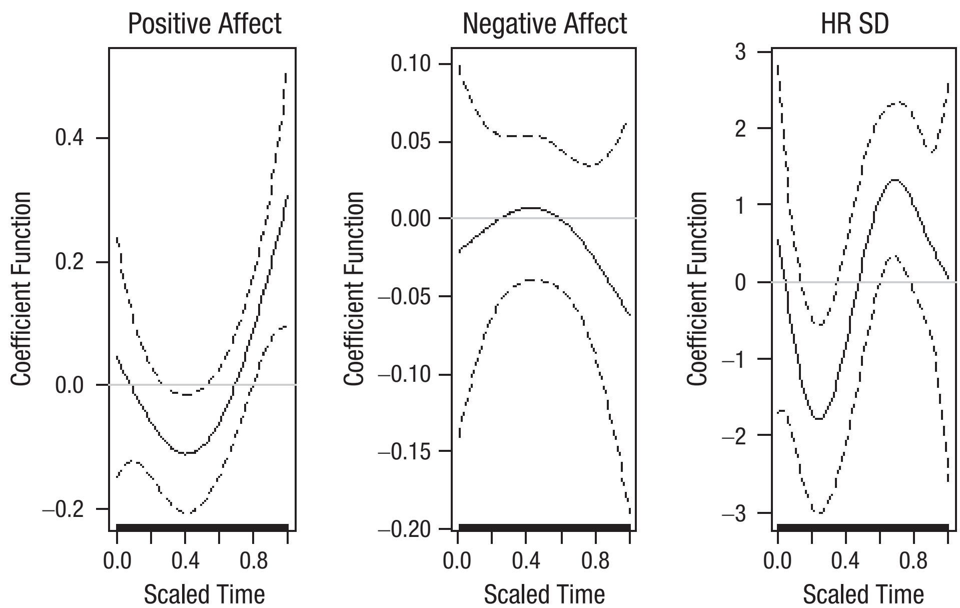


**Fig. 3.** Coefficient functions predicting participants' reports of (left) positive emotions, (center) negative emotions, and (right) change in depressive symptoms from within-sessions estimated valence trajectory. To facilitate interpretation, the change in depressive-symptom values were multiplied by –1. This adjustment indicates that the valence trajectory predicts a better outcome in terms of change in depressive symptoms.

we considered a variant of Model 4 with a more complex random-effects structure than Equation 5. This structure used orthogonal linear and quadratic functions $l(t), q(t)$ corresponding to an increasing linear function crossing zero at midsession and a parabola whose vertex (nadir) is at midsession. The model is expressed as

$$v_{ijk} = f_0(t_{ijk}) + (\beta + b_i)s_{ij} + \zeta_i + \eta_{ij} + w_{1i}l(t_{ijk}) + w_{2i}q(t_{ijk}) + \varepsilon_{ijk},$$

where $b_i$ is subject $i$'s random deviation from the mean slope with respect to session and $w_{1i}$, $w_{2i}$ are within-sessions linear and quadratic random effects; the remaining terms are as in Model 5. The purpose of this nonstandard random-effects structure was to extract the quadratic random effect $w_{2i}$ for each subject, a measure of the average "U-shapedness" of subject $i$'s within-sessions valence trajectory. The random effects $b_i$, $w_{1i}$, $w_{2i}$ were assumed to be jointly normal with mean of zero and not necessarily independent of each other, thus allowing us to assess whether $b_i$, which captures subject $i$'s between-sessions change in valence, was correlated with $w_{2i}$, the U-shapedness measure for subject $i$. In other words, this correlation measures whether patients who exhibited a stronger within-sessions pattern of decrease followed by an increase also exhibited a greater between-sessions increase in valence. This correlation between the within-sessions random quadratic effect $w_{2i}$ and the between-sessions random effect $b_i$ was found to be 0.77. Although this would ordinarily be considered a very high correlation, we note that correlations between different random effects in linear mixed-effect models are not accompanied by $p$ values and can be somewhat unstable. Therefore, we tested the positive correlation's stability by refitting the above model in 1,000 bootstrap samples (Tibshirani & Efron, 1993). In 50 of these models, the estimated variance of the between-sessions random effects $b_i$ was 0; hence, the correlation estimate was undefined. For 851 of the remaining 950 models (89.6%), the correlation was positive, yielding a two-sided $p$ value of .20 for the null hypothesis of zero correlation.

Finally, estimates of the coefficient function $\beta(t)$ in the scalar-on-function regression models are shown in Figure 3. For positive emotions, the valence function marginally predicted patients' prospective reports on their positive emotions ($F = 2.06$, $p = .07$). For negative emotions, no association was found for patients' reports on the NA scale. For PA scores, the pointwise confidence intervals around the coefficient function included zero throughout most of the sessions, but the coefficient function was pointwise significantly negative in the middle of the session and pointwise significantly positive at the end of the session. This suggests that lower valence (fewer positive and more negative emotions) in the middle of the session and higher valence (more positive and fewer negative emotions) at the end

**Table 2.** Estimated Between-Sessions Change in Valence According to Models 3 and 4

| | | Estimate | *SE* | *df* | *t* | *p* value |
|---|---|---|---|---|---|---|
| Model 3 | Session 4 | –0.251 | 1.307 | 1,5462.5 | –0.192 | .85 |
| | Session 8 | 1.389 | 1.293 | 1,5462.5 | 1.074 | .28 |
| | Session 12 | 2.111 | 1.377 | 1,5462.5 | 1.533 | .13 |
| | Session 16 | 3.684 | 1.404 | 1,5462.5 | 2.623 | .0087 |
| Model 4 | Slope | 0.256 | 0.084 | 1,5465.5 | 3.069 | .0022 |

Note: In both models, the dependent variable was valence, and the independent variable was session. In Model 3, the sessions were treated as categorical, and nonlinear between-sessions change was allowed. In Model 4, between-sessions change was linear and treated as continuous rather than categorical. In both models, the mean within-sessions trajectory had the same shape for each session.

of the session—that is, a U-shaped valence trajectory—predicts higher PA scores. This finding should be interpreted with caution in that the coefficient function was only marginally significant. For change in HRSD, the valence function is significantly predictive ($F = 2.72$, $p = .018$). As shown in Figure 3, when the valence was relatively low at the beginning of the session and increased toward the end of the session, this pattern predicted a decrease in depressive symptoms.

## Discussion

In the present study, we examined emotional avoidance, a primary mechanism underlying MDD, by delineating trajectories of change in emotions within and between sessions of psychological treatment for patients diagnosed with MDD. To the best of our knowledge, this is the first study to apply automatic coding of facial expressions and advanced statistical methods to explore these processes.

The results demonstrated that distinct and even opposite patterns of emotional change exist in the same data set depending on the temporal time frame of analysis. Consistent with the hypotheses, within a session, on average, the patients manifested a decrease in valence (more negative and fewer positive emotions) from the beginning to the middle of the session, followed by an increase in positive valence toward the end of the session. Alongside the temporary decrease in valence during the sessions, patients exhibited a consistent increase in valence throughout treatment.

Theoretically (e.g., S. C. Hayes et al., 1996), this pattern of change in emotions within and between sessions is consistent with the conceptualization of the emotional-avoidance model, which posits that over time, less recourse to emotional-avoidance strategies to regulate emotions may increase negative emotions in the short term but dampen the experience of negative emotions and facilitate the experience of positive ones in the long term. These results underscore the importance of overcoming emotional avoidance by showing that approaching negative emotions in the short term during treatment sessions may be beneficial in the long term for reducing negative emotions and increasing positive emotions. These changes are particularly crucial for patients diagnosed with MDD, who suffer from emotional disturbances and difficulties regulating emotions (Gross & Jazaieri, 2014).

This pattern of change in emotions closely resembles findings on the behavioral concept of habituation (Foa & Kozak, 1986; Mendolia & Kleck, 1993). “Habituation” broadly refers to a decline in response strength after repeated exposure to a stimulus. Previous studies addressing this concept in the context of negative emotions have suggested that repeated exposure to negative emotions results in a reduced response to such emotions through the process of habituation (Hunt, 1998; Low et al., 2008; Pascual-Leone et al., 2016). In these studies, participants were assigned to either emotional or control writing tasks following the induction of a distressing mood. Similar to the current findings, compared with the control group, participants in the study group manifested more negative emotional reactions immediately after the writing task but lesser reactions in subsequent assessments. The researchers argued that these emotional changes represent a process of activation and habituation of emotional responses across repeated presentations of a stressor (Hunt, 1998; Low et al., 2008). Thus, the current results lend weight to the recent suggestion that these principles of the habituation model can be applied to the treatment of MDD as well (A. M. Hayes et al., 2015).

In terms of the association between the U-shape valence trajectory and patients’ reports of their positive emotions, the results suggested that when patients experienced more negative emotions in the middle of the session and more positive emotions at the end of the sessions, they retrospectively reported having more positive emotions during the session. This finding may suggest that patients perceived positive emotions were more influenced by the process of emotion modification—in which negative emotions are transformed into positive ones—than by the actual expression of positive emotions (Berking et al., 2008). Alternatively, this finding may also imply that as a result of the regulation process, the way patients prospectively perceive these emotions undergoes a change, in addition to the felt changes in emotions. This finding also echoes the peak-end rule originally introduced by Kahneman et al. (1993), which states that people retrospectively evaluate subjective affective experiences by averaging the worst and the end of the experience. Similar findings were found in recent study in which acoustic features were extracted from couples' conversations to identify their emotions at the end of these sessions (Boateng et al., 2020). In the current study, patients' reports of negative emotions were not associated with changes in their expression of emotions during that session. If replicated in future studies, this finding may hint that one of the benefits of expressing negative emotions lies in making room for a greater perceived experience of positive emotions (Gleiser et al., 2008).

Finally, the results showed that when the valence was relatively low at the beginning of the session and increased toward the end of the session, this pattern predicted a decrease in depressive symptoms. This finding strengthens the theoretical assumption that expressing negative emotions during therapy sessions provides patients with an opportunity to modify and regulate their emotions, which may result in a reduction of depressive symptoms (Berking et al., 2008). Previous studies have relied on self-report measures administered after the session and found that improved emotion regulation was associated with a decrease in symptoms (e.g., Enrique et al., 2021). The current study is the first to use artificial-intelligence-based facial-expression analysis to study emotional dynamics, thus making it possible to track this dynamic as it unfolds during the session.

This study has several limitations that should be considered when interpreting the results. The primary limitation was the small sample size, which may have contributed to the weak association between in-session facial expressions and the patients' retrospective self-reports. In addition, the naturalistic setting of the study and the specific naturalistic type of data and measurement resulted in a significant amount of missing data. Limited resources also prevented us from analyzing all the sessions. However, a pilot study was conducted in which all the sessions of one patient were coded. The analyses revealed similar patterns of change in valence within and between sessions, which supports the generalizability of the results. Furthermore, the approach used to measure emotions, specifically valence, may be unable to disentangle positive from negative emotions and different type of negative emotions, making it difficult to determine the specific emotional experience. Finally, emotions were measured through only the single modality of facial expressions. Future studies should integrate additional indicators, such as the measurement of heart rate, to fully and more accurately capture the nonverbal expression of emotions (Sloan & Kring, 2007). If replicated in future studies, the present findings can be harnessed to advance theories of emotions in psychopathology and point to novel interventions that can guide therapists and improve mental health. Future studies could also explore how other treatment processes affect the patterns of emotions observed during the therapy sessions. For instance, studies could investigate how the therapeutic alliance influences the experience of positive or negative emotions at different points in the session, such as feeling more positive emotions at the start of the session when first engaging with the therapist or feeling more comfortable expressing negative emotions as the participants build trust with their therapist (Fisher et al., 2016).

Overall, the findings indicated that patients manifested a short-term decrease and a long-term increase in valence during psychological treatment for depression. In addition, when patients displayed a pattern of heightened negative facial expressions followed by an increase in positive emotions during therapy sessions, they also reported feeling more positive emotions during the session and had a subsequent reduction in depressive symptoms. These findings underscore the significance of allowing patients to confront their negative emotions as a means of enhancing their emotion-regulation capabilities and facilitating the amelioration of depressive symptoms.

**Transparency**

*Action Editor:* Pim Cuijpers
*Editor:* Jennifer L. Tackett
*Author Contribution(s)*

**Hadar Fisher:** Conceptualization; Data curation; Formal analysis; Project administration; Writing – original draft; Writing – review & editing.

**Philip T. Reiss:** Formal analysis; Funding acquisition; Supervision; Writing – review & editing.

**Dovrat Atias:** Data curation; Writing – review & editing.

**Michal Malka:** Validation; Writing – review & editing.

**Ben Shahar:** Conceptualization.

**Simone Shamay-Tsoory:** Conceptualization; Software; Writing – review & editing.

**Sigal Zilcha-Mano:** Conceptualization; Funding acquisition; Resources; Supervision; Writing – review & editing.

*Declaration of Conflicting Interests*

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

*Funding*

The work of Fisher and Reiss was supported by Israel Science Foundation under Grant 1076/19. The work of Fisher was also supported by the Data Science Research Center, University of Haifa. The study was supported by Israel Science Foundation under Grant 186/15 and by the American Psychological Foundation (APF) Walter Katkovsky Research Grant.

*Open Practices*

This article has received badge for Open Materials. More information about the Open Practices badges can be found at https://www.psychologicalscience.org/publications/badges

## ORCID iDs

Hadar Fisher https://orcid.org/0000-0002-3421-3966
Simone Shamay-Tsoory https://orcid.org/0000-0003-1088-5212
Sigal Zilcha-Mano https://orcid.org/0000-0002-5645-4429

## Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/21677026231195793

## Note

1. Note that the estimates provided by AFFDEX software may be considerably influenced by the detection of smiles because they are commonly and readily detected by the software. To examine whether the results were affected by the detection of smiles, we conducted additional post hoc analyses solely for negative emotions. The anticipated pattern (an inverted U-shape) was observed. For a complete description of the analyses and findings, see the Supplemental Material available online.

## References

Aafjes-van Doorn, K., & Barber, J. P. (2017). Systematic review of in-session affect experience in cognitive behavioral therapy for depression. *Cognitive Therapy and Research*, *41*, 807–828.

Alghowinem, S., Goecke, R., Cohn, J. F., Wagner, M., Parker, G., & Breakspear, M. (2015). Cross-cultural detection of depression from nonverbal behaviour. In *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)* (Vol. 1, pp. 1–8). IEEE.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.).

Arango, I., Miranda, E., Ferrer, J. C. S., Fresán, A., Ortega, M. A. R., Vargas, A. N., Barragán, S., Villanueva-Valle, J., Munoz-Delgado, J., & Robles, R. (2019). Changes in facial emotion expression during a psychotherapeutic intervention for patients with borderline personality disorder. *Journal of Psychiatric Research*, *114*, 126–132.

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology*, *8*, Article 456. https://doi.org/10.3389/fpsyg.2017.00456

Barrett, L. F. (1998). Discrete emotions or dimensions? The role of valence focus and arousal focus. *Cognition & Emotion*, *12*(4), 579–599.

Beringer, M., Spohn, F., Hildebrandt, A., Wacker, J., & Recio, G. (2019). Reliability and validity of machine vision for the assessment of facial expressions. *Cognitive Systems Research*, *56*, 119–132.

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.

Bland, J. M., & Altman, D. G. (1995). Statistics notes: Calculating correlation coefficients with repeated observations: Part 1—Correlation within subjects. *BMJ*, *310*, Article 446. https://doi.org/10.1136/bmj.310.6977.446

Boateng, G., Sels, L., Kuppens, P., Hilpert, P., & Kowatsch, T. (2020). Speech emotion recognition among couples using the peak-end rule and transfer learning. In *Companion Publication of the 2020 International Conference on Multimodal Interaction* (pp. 17–21). Association for Computing Machinery.

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.

Campbell-Sills, L., Barlow, D. H., Brown, T. A., & Hofmann, S. G. (2006). Acceptability and suppression of negative emotion in anxiety and mood disorders. *Emotion*, *6*(4), 587–597.

Cootes, T. F., Edwards, G. J., & Taylor, C. J. (2001). Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*(6), 681–685.

Cribb, G., Moulds, M. L., & Carter, S. (2006). Rumination and experiential avoidance in depression. *Behaviour Change*, *23*(3), 165–176.

Cuijpers, P., Karyotaki, E., Weitz, E., Andersson, G., Hollon, S. D., & van Straten, A. (2014). The effects of psychotherapies for major depression in adults on remission, recovery and improvement: A meta-analysis. *Journal of Affective Disorders*, *159*, 118–126.

Diener, M. J., Hilsenroth, M. J., & Weinberger, J. (2007). Therapist affect focus and patient outcomes in psychodynamic psychotherapy: A meta-analysis. *American Journal of Psychiatry*, *164*(6), 936–941.

Dziak, J. J., Coffman, D. L., Reimherr, M., Petrovich, J., Li, R., Shiffman, S., & Shiyko, M. P. (2019). Scalar-on-function regression for predicting distal outcomes from intensively gathered longitudinal data: Interpretability for applied scientists. *Statistics Surveys*, *13*, 150–180.

Ekman, P., & Rosenberg, E. (2005). *What the face reveals* (2nd ed.). Oxford University Press.

Ellgring, H. F. (1989). *Nonverbal communication in depression*. Cambridge University Press.

Enrique, A., Eilert, N., Wogan, R., Earley, C., Duffy, D., Palacios, J., Timulak, L., & Richards, D. (2021). Are changes in beliefs about rumination and in emotion regulation skills mediators of the effects of internet-delivered cognitive-behavioral therapy for depression and anxiety? Results from a randomized controlled trial. *Cognitive Therapy and Research*, *45*, 805–816. https://doi.org/10.1007/s10608-020-10200-6

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.

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.

Fisher, H., Rafaeli, E., Bar-Kalifa, E., Barber, J. P., Solomonov, N., Peri, T., & Atzil-Slonim, D. (2020). Therapists' interventions as a predictor of clients' emotional experience, self-understanding, and treatment outcomes. *Journal of Counseling Psychology*, *67*(1), 66–78.

Fleiss, J. L., Levin, B., & Paik, M. C. (2013). *Statistical methods for rates and proportions*. John Wiley & Sons.

Flynn, M., Effraimidis, D., Angelopoulou, A., Kapetanios, E., Williams, D., Hemanth, J., & Towell, T. (2020). Assessing the effectiveness of automated emotion recognition in adults and children for clinical investigation. *Frontiers in Human Neuroscience*, *14*, Article 70. https://doi.org/10.3389/fnhum.2020.00070

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: exposure to corrective information. *Psychological Bulletin*, *99*(1), 20–35.

Forbes, C. N., Tull, M. T., Xie, H., Christ, N. M., Brickman, K., Mattin, M., & Wang, X. (2020). Emotional avoidance and social support interact to predict depression symptom severity one year after traumatic exposure. *Psychiatry Research*, *284*, Article 112746. https://doi.org/10.1016/j.psychres.2020.112746

Friedrich, M. J. (2017). Depression is the leading cause of disability around the world. *Journal of the American Medical Association*, *317*(15), 1517. https://doi.org/10.1001/jama.2017.3826

Girard, J. M., & Cohn, J. F. (2015). Automated audiovisual depression analysis. *Current Opinion in Psychology*, *4*, 75–79.

Girard, J. M., Cohn, J. F., Mahoor, M. H., Mavadati, S., & Rosenwald, D. P. (2013). Social risk and depression: Evidence from manual and automatic facial expression analysis. In *10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)* (pp. 1–8). IEEE.

Girard, J. M., Cohn, J. F., Mahoor, M. H., Mavadati, S. M., Hammal, Z., & Rosenwald, D. P. (2014). Nonverbal social withdrawal in depression: Evidence from manual and automatic analyses. *Image and Vision Computing*, *32*(10), 641–647.

Gleiser, K., Ford, J. D., & Fosha, D. (2008). Contrasting exposure and experiential therapies for complex posttraumatic stress disorder. *Psychotherapy: Theory, Research, Practice, Training*, *45*(3), 340–360.

Goldsmith, J., Scheipl, F., Huang, L., Wrobel, J., Di, C., Gellar, J., Harezlak, J., McLean, M. W., Swihart, B., Xiao, L., Crainiceanu, C., & Reiss, P. T. (2022). *refund: Regression with functional data* (R package Version 0.1-26). https://cran.r-project.org/web/packages/refund/refund.pdf

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.

Greenberg, L., & Safran, J. (1990). Emotional change processes in psychotherapy. In R. Plutchik & H. Kellerman (Eds.), *Emotion: Theory, research, and experience* (Vol. 5, pp. 59–85). Academic Press.

Gross, J. J. (2013). Emotion regulation: Taking stock and moving forward. *Emotion*, *13*(3), 359–365.

Gross, J. J., & Jazaieri, H. (2014). Emotion, emotion regulation, and psychopathology: An affective science perspective. *Clinical Psychological Science*, *2*(4), 387–401.

Gunes, H., & Schuller, B. (2013). Categorical and dimensional affect analysis in continuous input: Current trends and future directions. *Image and Vision Computing*, *31*(2), 120–136.

Gupta, T., Haase, C. M., Strauss, G. P., Cohen, A. S., Ricard, J. R., & Mittal, V. A. (2022). Alterations in facial expressions of emotion: Determining the promise of ultrathin slicing approaches and comparing human and automated coding methods in psychosis risk. *Emotion*, *22*(4), 714–724.

Habes, I., Krall, S. C., Johnston, S. J., Yuen, K. S. L., Healy, D., Goebel, R., Sorger, B., & Linden, D. E. J. (2013). Pattern classification of valence in depression. *NeuroImage: Clinical*, *2*, 675–683. https://doi.org/10.1016/j.nicl.2013.05.001

Haines, N., Bell, Z., Crowell, S., Hahn, H., Kamara, D., McDonough-Caplan, H., Shader, T., & Beauchaine, T. P. (2019). Using automated computer vision and machine learning to code facial expressions of affect and arousal: Implications for emotion dysregulation research. *Development and Psychopathology*, *31*(3), 871–886.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, *23*(1), 56–62.

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Social and Clinical Psychology*, *6*(4), 278–296.

Harley, J. M. (2016). Measuring emotions: A survey of cutting edge methodologies used in computer-based learning environment research. In S. Y. Tettegah & M. Gartmeier (Eds.), *Emotions, technology, design, and learning* (pp. 89–114). Elsevier.

Hastie, T., & Tibshirani, R. (1993). Varying-coefficient models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *55*(4), 757–779.

Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But. . . . *Communication Methods and Measures*, *14*(1), 1–24.

Hayes, A. M., Yasinski, C., Barnes, J. B., & Bockting, C. L. (2015). Network destabilization and transition in depression: New methods for studying the dynamics of therapeutic change. *Clinical Psychology Review*, *41*, 27–39.

Hayes, S. C., Wilson, K. G., Gifford, E. V., Follette, V. M., & Strosahl, K. (1996). Experiential avoidance and behavioral disorders: A functional dimensional approach to diagnosis and treatment. *Journal of Consulting and Clinical Psychology*, *64*(6), 1152–1168.

Höfling, T. T. A., Alpers, G. W., Gerdes, A. B., & Föhl, U. (2021). Automatic facial coding versus electromyography of mimicked, passive, and inhibited facial response to emotional faces. *Cognition and Emotion*, *35*(5), 874–889.

Hunt, M. G. (1998). The only way out is through: Emotional processing and recovery after a depressing life event. *Behaviour Research and Therapy*, *36*(4), 361–384.

iMotions. (2022). *iMotions biometric research platform* (Version 9.3) [Computer software]. https://imotions.com/

Joormann, J., & Stanton, C. H. (2016). Examining emotion regulation in depression: A review and future directions. *Behaviour Research and Therapy*, *86*, 35–49.

Joormann, J., & Vanderlind, W. M. (2014). Emotion regulation in depression: The role of biased cognition and reduced cognitive control. *Clinical Psychological Science*, *2*(4), 402–421.

Kahn, J. H., & Garrison, A. M. (2009). Emotional self-disclosure and emotional avoidance: Relations with symptoms of depression and anxiety. *Journal of Counseling Psychology*, *56*(4), 573–584.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science*, *4*(6), 401–405.

Kim, J., Weber, C. E., Gao, C., Schulteis, S., Wedell, D. H., & Shinkareva, S. V. (2020). A study in affect: Predicting valence from fMRI data. *Neuropsychologia*, *143*, Article 107473. https://doi.org/10.1016/j.neuropsychologia.2020.107473

Ko, B. C. (2018). A brief review of facial emotion recognition based on visual information. *Sensors*, *18*(2), 401.

Krumhuber, E. G., Küster, D., Namba, S., Shah, D., & Calvo, M. G. (2021a). Emotion recognition from posed and spontaneous dynamic expressions: Human observers versus machine analysis. *Emotion*, *21*(2), 447–451.

Krumhuber, E. G., Küster, D., Namba, S., & Skora, L. (2021b). Human and machine validation of 14 databases of dynamic facial expressions. *Behavior Research Methods*, *53*(2), 686–701.

Kulke, L., Feyerabend, D., & Schacht, A. (2020). A comparison of the Affectiva iMotions Facial Expression Analysis Software with EMG for identifying facial expressions of emotion. *Frontiers in Psychology*, *11*, Article 329. https://doi.org/10.3389/fpsyg.2020.00329

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin*, *139*(4), 917–940.

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods*, *9*(2), 202–220.

Lewinski, P., Den Uyl, T. M., & Butler, C. (2014). Automated facial coding: Validation of basic emotions and FACS AUs in FaceReader. *Journal of Neuroscience, Psychology, and Economics*, *7*(4), 227–236.

Low, C. A., Stanton, A. L., & Bower, J. E. (2008). Effects of acceptance-oriented versus evaluative emotional processing on heart rate recovery and habituation. *Emotion*, *8*(3), 419–424.

Lucas, G. M., Gratch, J., Scherer, S., Boberg, J., & Stratou, G. (2015). Towards an affective interface for assessment of psychological distress. In *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)* (pp. 539–545). IEEE.

Magdin, M., Benko, Ľ., & Koprda, Š. (2019). A case study of facial emotion classification using Affdex. *Sensors*, *19*(9), Article 2140. https://doi.org/10.3390/s19092140

McDuff, D., Mahmoud, A., Mavadati, M., Amr, M., Turcot, J., & Kaliouby, R. (2016). AFFDEX SDK: A cross-platform real-time multi-face expression recognition toolkit. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 3723–3726). Association for Computing Machinery.

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *23*(3), 412–433.

Mellick, W. H., Mills, J. A., Kroska, E. B., Calarge, C. A., Sharp, C., & Dindo, L. N. (2019). Experiential avoidance predicts persistence of major depressive disorder and generalized anxiety disorder in late adolescence. *The Journal of Clinical Psychiatry*, *80*(6), Article 18m12265. https://doi.org/10.4088/JCP.18m12265

Mendolia, M., & Kleck, R. E. (1993). Effects of talking about a stressful event on arousal: Does what we talk about make a difference? *Journal of Personality and Social Psychology*, *64*(2), 283–292.

Pampouchidou, A., Marias, K., Tsiknakis, M., Simos, P., Yang, F., Lemaître, G., & Meriaudeau, F. (2016). Video-based depression detection using local curvelet binary patterns in pairwise orthogonal planes. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 3835–3838). IEEE.

Pascual-Leone, A., Yeryomenko, N., Morrison, O.-P., Arnold, R., & Kramer, U. (2016). Does feeling bad, lead to feeling good? Arousal patterns during expressive writing. *Review of General Psychology*, *20*(3), 336–347.

Peluso, P. R., & Freund, R. R. (2018). Therapist and client emotional expression and psychotherapy outcomes: A meta-analysis. *Psychotherapy*, *55*(4), 461–472.

Power, M. J., & Tarsia, M. (2007). Basic and complex emotions in depression and anxiety. *Clinical Psychology & Psychotherapy*, *14*(1), 19–31.

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Reiss, P. T., Goldsmith, J., Shang, H. L., & Ogden, R. T. (2017). Methods for scalar-on-function regression. *International Statistical Review*, *85*(2), 228–249.

Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric regression* (Issue 12). Cambridge University Press.

Sandbach, G., Zafeiriou, S., Pantic, M., & Yin, J. (2012). Static and dynamic 3D facial expression recognition: A comprehensive survey. *Image and Vision Computing*, *30*, 683–697.

Scherer, S., Stratou, G., Lucas, G., Mahmoud, M., Boberg, J., Gratch, J., & Morency, L. P. (2014). Automatic audiovisual behavior descriptors for psychological disorder analysis. *Image and Vision Computing*, *32*(10), 648–658.

Scherer, S., Stratou, G., & Morency, L. P. (2013). Audiovisual behavior descriptors for depression assessment In *Proceedings of the 15th ACM on International conference on multimodal interaction* (pp. 135–140). ACM.

Schneider, F., Heimann, H., Himer, W., Huss, D., Mattes, R., & Adam, B. (1990). Computer-based analysis of facial action in schizophrenic and depressed patients. *European Archives of Psychiatry and Clinical Neuroscience*, *240*(2), 67–76.

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. *Journal of Clinical Psychiatry*, *59*(20), 22–33.

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.

Stöckli, S., Schulte-Mecklenbeck, M., Borer, S., & Samson, A. C. (2018). Facial expression analysis with AFFDEX and FACET: A validation study. *Behavior Research Methods*, *50*, 1446–1460.

Stratou, G., Scherer, S., Gratch, J., & Morency, L. P. (2015). Automatic nonverbal behavior indicators of depression and PTSD: The effect of gender. *Journal on Multimodal User Interfaces*, *9*(1), 17–29.

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

Ten Hove, D., Jorgensen, T. D., & van der Ark, L. A. (2022). Interrater reliability for multilevel data: A generalizability theory approach. *Psychological Methods*, *27*, 650–666.

Thoma, N. C., & McKay, D. (2014). Introduction. In N. C. Thoma & D. McKay (Eds.), *Working with emotion in cognitive behavioral therapy: Techniques for clinical practice* (pp. 1–8). The Guilford Press.

Tibshirani, R. J., & Efron, B. (1993). An introduction to the bootstrap. *Monographs on Statistics and Applied Probability*, *57*, 1–436.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

Tull, M. T., Gratz, K. L., Salters, K., & Roemer, L. (2004). The role of experiential avoidance in posttraumatic stress symptoms and symptoms of depression, anxiety, and somatization. *The Journal of Nervous and Mental Disease*, *192*(11), 754–761.

Watson, D., Clark, L. A., & Stasik, S. M. (2011). Emotions and the emotional disorders: A quantitative hierarchical perspective. *International Journal of Clinical and Health Psychology*, *11*(3), 429–442.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070.

Waxer, P. H. (1974). Therapist training in nonverbal communication: I. Nonverbal cues for depression. *Journal of Clinical Psychology*, *30*, 215–218.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *73*(1), 3–36.

Wood, S. N. (2013). On p-values for smooth components of an extended generalized additive model. *Biometrika*, *100*(1), 221–228.

Wood, S. N. (2017). P-splines with derivative based penalties and tensor product smoothing of unevenly distributed data. *Statistics and Computing*, *27*(4), 985–989.

Wood, S. N., Pya, N., & Säfken, B. (2016). Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association*, *111*(516), 1548–1563.

Wood, S. N., & Scheipl, F. (2020). *gamm4: Generalized Additive Mixed Models using "mgcv" and "lme4"* (R package Version 0.2-6). https://cran.r-project.org/package=gamm4

Yitzhak, N., Giladi, N., Gurevich, T., Messinger, D. S., Prince, E. B., Martin, K., & Aviezer, H. (2017). Gently does it: Humans outperform a software classifier in recognizing subtle, nonstereotypical facial expressions. *Emotion*, *17*, 1187–1198. https://doi.org/10.1037/emo0000287

Zilcha-Mano, S., Shahar, B., Fisher, H., Dolev-Amit, T., Greenberg, L. S., & Barber, J. P. (2021). Investigating patient-specific mechanisms of change in SET vs. EFT for depression: Study protocol for a mechanistic randomized controlled trial. *BMC Psychiatry*, *21*(1), Article 287. https://doi.org/10.1186/s12888-021-03279-y1–11.