The effects of psychotherapy for major depressive disorder on daily mood and functioning: A longitudinal experience sampling study

By: <u>Kari M. Eddington</u>, Chris J. Burgin, <u>Paul J. Silvia</u>, Niloofar Fallah, Catherine Majestic, and <u>Thomas J. Kwapil</u>

Eddington, K.M, Burgin, C.J., Silvia, P.J., Fallah, N., Majestic, C., & Kwapil, T.R. (2017). The effects of psychotherapy for major depressive disorder on daily mood and functioning: A longitudinal experience sampling study. *Cognitive Therapy and Research*, *41(2)*, 266-277. PubMed Central PMCID: PMC5451128.

This is a post-peer-review, pre-copyedit version of an article published in *Cognitive Therapy and Research*. The final authenticated version is available online at: <u>http://dx.doi.org/10.1007/s10608-016-9816-7</u>

***© 2016 Springer Science+Business Media New York. Reprinted with permission. No further reproduction is authorized without written permission from Springer US. This version of the document is not the version of record. This document is subject to <u>Springer</u> terms of re-use. ***

Abstract:

Experience sampling methodology (ESM) was used in a randomized controlled trial of shortterm therapy to examine changes in daily affect and reactivity to daily event appraisals among depressed patients. Fifty-five depressed adults (mean age 37 years, 80 % female) were randomly assigned to one of two therapy conditions. Using an interactive voice response system, participants rated activities and emotional functioning eight times per day for 7 days. Twentynine participants completed treatment and repeated ESM at post-treatment. Broad improvements in mood, cognition, and physical functioning were similar across treatment conditions, with the largest improvements for markers of positive affect. Participants demonstrated increased resilience, i.e., diminished reactivity to stressors, at post-treatment. Changes in reactivity to positive daily situations were minimal. Findings underscore the utility of ESM in psychotherapy research and the importance of including measures of both positive and negative affect and experiences.

Keywords: Major depressive disorder | Emotions | Emotional stress | Psychotherapy | Treatment outcome | Cognitive behavior therapy

Article:

Introduction

Evaluating the efficacy of psychological interventions for depression typically involves assessments before and after treatment, often with measures inquiring about symptoms over a retrospective period of several days or weeks. For example, the Beck Depression Inventory (Beck et al. 1996) and Patient Health Questionnaire 9-item depression scale (Kroenke et al. 2001) ask about symptoms over the past 2 weeks; the Center for Epidemiological Studies Depression scale (Radloff 1977) covers the past week. The respondents' ratings should reflect "typical" symptom levels; however, the accuracy of recalled information has been called into question (Schrader et al. 1990).

Experience sampling methodology (ESM) refers to data collection methods in which respondents are assessed during daily life multiple times over days, weeks, or months (Conner et al. 2009). Studies using ESM, which asks about in-the-moment experiences, alongside traditional self-reports have shown biases in the retrospective reporting of emotions, particularly when people are asked to produce "averages" over days or weeks (Ben-Zeev and Young 2010; Ben-Zeev et al. 2009; Parkinson et al. 1995). Furthermore, with ESM, researchers can evaluate within-person variability over shorter time windows. Another important advantage of ESM is that data are collected in the participant's natural environment, allowing for fine-grained examination of contextual factors (Ebner-Priemer and Trull 2009).

One such contextual factor, daily positive and negative events, has been examined in several studies using ESM. Results on reactivity to stressors have varied, with some showing no differences between depressed and control groups (Bylsma et al. 2011) and others showing exaggerated negative affect (NA) in clinical depression (Myin-Germeys et al. 2003), dysphoria, and anxiety (Nezlek and Gable 2001; South and Miller 2014). One study found that the NA response to stress in depression, while blunted in comparison to controls, was more persistent over time (Peeters et al. 2003). Regarding positive events, some ESM studies have reported a "mood brightening" effect [increase in positive affect (PA) and decrease in NA with positive events] among depressed patients. Bylsma et al. (2011) examined event reactivity and reported that, relative to controls, patients had a greater decrease in NA in response to positive events, with no differences in PA. Using a slightly more severe sample, Peeters and colleagues found larger PA increases and smaller NA decreases in patients (Peeters et al. 2003).

Regarding treatment effects, ESM is becoming increasingly common in studies for some conditions such as substance abuse (Moore et al. 2014; Tidey et al. 2008) but has been used infrequently with depression. Barge-Schaapveld and colleagues used ESM to evaluate the efficacy of antidepressants (Barge-Schaapveld and Nicolson 2002; Barge-Schaapveld et al. 1995; Wichers et al. 2009) and found that depressed patients had higher NA responses to stress than controls at baseline, a difference that decreased after 6 weeks of pharmacotherapy. Patients did not differ from controls on PA responses to pleasant events at baseline, but PA responses increased with treatment (Wichers et al. 2009).

A handful of studies comparing different forms of treatment (without nondepressed controls or a no-treatment group) have been conducted with ESM. Peeters et al. (2010) used 6 days of pretreatment ESM in an uncontrolled depression treatment study (pharmacotherapy and supportive therapy combined). They found that lower NA in response to negative events in daily life predicted poorer outcomes, interpreted as indicating that a generalized blunting of affective reactivity is associated with poorer prognosis. However, their study did not include a post-treatment ESM assessment, so pre/post changes could not be assessed. Another study (Forbes et al. 2012) used ESM throughout an 8-week treatment in a sample of children and adolescents with depression and anxiety. Treatment conditions were not randomly assigned and included cognitive-behavioral therapy (CBT), pharmacotherapy, or both. They found that ratings of emotion assessed using ESM predicted response to treatment above and beyond retrospective self-report measures. A third study using ESM examined mindfulness-based cognitive therapy (MBCT) as an adjunct to treatment-as-usual in patients with residual depressive. Therapy was associated with increases in PA experiences in general and in response to pleasant events (Geschwind et al. 2011).

The results from these studies suggest that momentary assessments provide a unique contribution to our understanding of treatment response. They are becoming an increasingly prominent therapeutic tool for clinicians (Naylor et al. 2008; Rose et al. 2012); for example, a recent study used ESM-based feedback as an adjunct to treatment as usual (medications and/or psychotherapy) for depression (Hartmann et al. 2015). However, there have been no studies to our knowledge using ESM to evaluate outcomes in the context of a randomized controlled psychotherapy trial. In the current study, depressed patients were randomly assigned to CBT or self-system therapy (SST). The treatments differed primarily in their proposed mechanisms of action: whereas CBT focuses on restructuring dysfunctional cognitions, SST targets problems with self-regulation—the ability to set, monitor, and successfully pursue personal goals (Vieth et al. 2003). Previous studies (Eddington et al. 2015; Strauman et al. 2006) showed that the two therapies were equally effective in reducing depressive symptoms using retrospective self-report measures. However, both studies showed that SST was more effective for participants with pre-treatment deficits in self-regulation.

Participants in the Eddington et al. clinical trial completed a week of intensive experience sampling at baseline and during the week preceding the final therapy session. Using a phone-based interactive voice response (IVR) system, participants responded to questions in five domains: mood, cognition, social activity, physical functioning, and activities. Because the two treatments share a common goal of improving emotional functioning and coping (albeit through different means), both should show increases in PA and decreases in NA. Because the ESM items assessed broad domains of functioning (for example, the cognition items tapped into broad aspects of cognition like trouble concentrating rather than treatment-specific ones like cognitive distortions), we did not expect condition-specific effects.

A secondary aim was to examine the extent to which emotional reactivity to daily stressors and positive events changes with treatment. Findings regarding affective blunting and mood brightening notwithstanding, we expected that patients should show decreased NA in response to stress and increased PA in response to positive events. This hypothesis was guided by what the treatments are supposedly doing (improving skills for coping with stress and enhancing focus on pleasant events) and by the handful of studies showing this pattern of change with pharmacotherapy (e.g., Barge-Schaapveld et al. 1995; Wichers et al. 2009).

Methods

Participants

Participants were 55 (11 male, 44 female; mean age = 36.7 years, SD = 13.3) adults recruited via advertisements. Study procedures were approved by an Institutional Review Board and informed

consent was obtained from all participants. Demographic characteristics are shown in Table 1; for a more detailed review of recruitment procedures and participant flow see Eddington et al. (2015) and Fig. 1. Inclusion criteria included primary diagnosis of major depressive disorder (MDD) or dysthymia, BDI-II score above 14, no antidepressant medication use for 4 months prior to the study, no active substance use problems for 6 months prior, no history of mania, no history of psychotic symptoms, no diagnosis of antisocial or borderline personality disorders, and no active suicidal intent or immediate threats of self-harm.

Age	M = 36.73 $SD = 13.17$
Cov	$\frac{11}{11} \text{ male } 44 \text{ female}$
	11 maie, 44 female
Highest level of education	
High school or below	18.2 %
Some college or college degree	65.4 %
Some graduate school or advanced degree	16.4 %
Employment status	65.5 % employed
Marital status	
Married	23.6 %
Divorced	20.0 %
Separated	5.5 %
Never married	50.9 %
Diagnostic status	
Primary MDD	96.4 %
Primary dysthymia	3.6 %
Secondary anxiety disorder ^a	72.7 %
Treatment history	
Past therapy	85.5 %
Past medications	36.4 %
Past hospitalization	30.9 %
BDI-II score at baseline (M, SD)	M = 34.71, SD = 8.14
BAI score at baseline (M, SD)	M = 23.24, SD = 14.86

Table 1. Participant descriptive information

^a This category also includes PTSD

Of 100 screened, 56 met inclusion criteria¹ and were assigned to either SST (n = 27) or CBT (n = 29) using restricted randomization. One participant who was randomized to SST but completed only one ESM survey was excluded. Of the 55 participants with baseline ESM data, 29 (15 in SST and 14 in CBT) also completed the post-treatment ESM. Six participants did not initiate any treatment, one participant completed treatment but did not complete post-treatment ESM, and the remaining 19 (34 %) completed at least one therapy session and therefore are considered drop-outs. There were no significant differences on baseline BAI and BDI-II between those who completed or prematurely dropped out of treatment (p's > .05). Of the 15 ESM items used in the current analyses, only the item, "My emotions feel out of control right now" significantly differed at baseline between the two groups using within-person means,

¹ Of the 44 exclusions, 38 % had an exclusionary diagnosis (e.g., bipolar); 35 % had subclinical depressive symptoms; 27 % had other exclusionary concerns (e.g., taking antidepressants, unable to follow through with study procedures).

t(53) = 4.61, p < .05. Participants who dropped out of treatment (M = 2.95, SD = 1.26) rated that item somewhat higher on average than completers (M = 2.33, SD = 0.98).



Figure 1. CONSORT participant enrollment and flow diagram

Measures

Semi-structured Clinical Interview for DSM-IV-TR: Research Version

The SCID-I (research version) and SCID-II are semi-structured diagnostic interviews for DSM-IV-TR (American Psychiatric Association 2000) Axis I and Axis II disorders, respectively. From the SCID-I, study diagnosticians administered the overview and screening sections, mood modules, anxiety modules, and psychotic symptoms; from the SCID-II, the antisocial and borderline personality modules were used. Additional modules (e.g., substance abuse and dependence, eating disorders) were used as needed based on the participant's responses to the screening items.

Beck Depression Inventory-II

The BDI-II (Beck et al. 1996) is a self-report measure of depressive symptom severity during the past 2 weeks that contains 21 items, each rated on a 4-point scale (0–3). The BDI-II has excellent psychometric properties (Dozois et al. 1998), and internal consistency in this study was good ($\alpha = .82$).

Table 2. Experience sampling questionnaire items

Please rate the following statements based upon what you were thinking, feeling, and doing at the time of the call from 1 <i>not at all</i> to 7 <i>very much</i> with four indicating <i>somewhat</i> . It is important that you describe how you were
feeling at the moment just before you received the phone call
Cognition
My thoughts are pleasant right now
I have trouble concentrating right now
I am ruminating right now
Mood
I feel happy right now
I feel confident right now
I feel satisfied right now
I feel good about myself right now
I feel like things are going well right now
I feel sad right now
My emotions feel out of control right now
I feel anxious right now
I feel bad about myself right now
I feel guilty right now
I feel irritable right now
I feel overwhelmed right now
Activities
I like what I am doing right now
I am successful in my current activity
I would prefer to do something else right now
I am not doing what I should be doing right now
Right now there is nothing I want to do
Physical functioning
I feel tired right now
I feel energetic right now
I don't feel well right now
Social functioning
Are you alone at this time?
If with others:
I feel close to this person (these people)
Right now I would prefer to be alone
If alone:
I am alone right now because people do not want to be with me
Right now I would prefer to be with other people
Appraisals of situation
My current situation is stressful
My current situation is positive

People responded to the item "Are you alone at this time?" using a binary *yes/no* scale, which was then coded as 0 = with others, 1 = alone. For all item categories (except for the Social Functioning and Appraisals of Situation items), the items appeared in a random within-category order at each call

Beck Anxiety Inventory

The BAI (Beck et al. 1988) is a self-report measure of anxiety symptom severity during the past 2 weeks. The scale consists of 21 items, each rated on a 4-point scale (0–3) scale. The BAI has good psychometric properties (Fydrich et al. 1992); in this study, internal consistency was very good ($\alpha = .95$).

Experience Sampling Methodology Items

An interactive voice response system (IVR) was used for ESM data collection. IVR involves the use of a host computer that is programmed to automatically call participants' cell phones and (using pre-recorded voice prompts) asks the participants questions to which they respond using their phone keypad. Participants were prompted via cell phone to complete 31 items assessing aspects of cognition, positive and negative mood, and experience of contextual factors such as activities, physical functioning, social functioning, and stressful situations at the time of the call (Table 2). Items were not constructed in an attempt to measure unitary constructs. Recognizing the importance of keeping each survey extremely brief in order to minimize participant burden, we tried to avoid redundancy among the items. Items in the "mood" category were modeled after items found in many widely-used measures of current mood, such as the Profile of Mood States (Shacham 1983) and Positive and Negative Affect Schedule (Watson et al. 1988). Participants entered responses using their mobile phone digital keypad.

Procedures

During intake participants provided informed consent in compliance with Institutional Review Board standards, completed the BDI-II and BAI, and (if the BDI-II score was above 14) completed relevant portions of the SCID. Diagnostic interviews were conducted by clinical psychology doctoral students who received training on the administration of the SCID-I and SCID-II. Following completion of the diagnostic interviews, participants were informed of the frequency and duration of ESM calls were given a brief overview of how to use the IVR system (Telesage; www.telesage.com) with their own phone or one provided to them. Participants designated a 12-h window that they preferred to receive calls.

The day following intake began their full week of baseline assessment. Baseline and posttreatment intensive ESM data collection employed random-time sampling to signal participants eight times each day for 1 week. The eight calls were randomized within 90 min segments equally separated within each 12-h block. If a participant missed a call, he or she could call the system back within a restricted 10-min period, and responses were all date- and time-stamped. Items and timing parameters were identical at pre- and post-treatment. Upon completion of the baseline ESM, participants were randomly assigned treatment (treatment assignments were unknown to the participants during the baseline ESM week). Details about the two therapy conditions can be found in Eddington et al. (2015). Briefly, both skills-based treatments involved 16 sessions oriented toward current problems and concerns. Adherence to treatment protocols was assessed through blind coding of therapy session recordings and the two conditions differed in the expected directions (CBT therapists used significantly more CBT strategies; SST therapists used significantly more SST strategies).

Results

Analytic Approach and ESM Compliance

Given the nested structure of the data, our primary analytic approach was multilevel modeling. We estimated the multilevel models with HLM 7, using restricted maximum likelihood with robust standard errors. All items were modeled as continuous except for the item "Are you alone at this time?", which was modeled as binary. Unless noted otherwise, all effects are unstandardized.

ESM items were analyzed separately, allowing for a more fine-grained analysis of specific changes from pre- to post-treatment. The use (and analysis) of single face valid items to assess straightforward variables is common practice in ESM research (Barrantes-Vidal et al. 2013; Depp et al. 2012; Fisher and To 2012) For example, a single item was shown to be a valid indicator of fatigue when compared with a multi-item scale (van Hooff et al. 2007). Within- and between-subjects correlations among the ESM items are available upon request from the corresponding author.

Missing data in experience sampling work is primarily "signal wise": people respond to some but not all calls (Silvia et al. 2013), so people vary in the number of calls they completed. The response rate was good and comparable to past studies with general samples (Burgin et al. 2012). At baseline, people completed on average 72.23 % of the calls (M = 40.45 calls completed, SD = 14.34, Min/Max = 6, 58); for those who completed a full course of treatment, people completed on average 63.66 % of the calls at post-treatment (M = 35.65 calls completed, SD = 14.49, Min/Max = 3, 56); a difference that was significant t(28) = 3.50, p < .01. The multilevel models were specified so that the effective sample size was the number of people who completed the first ESM week, so all participants were retained in the analysis regardless of whether they completed a full course of treatment.

Change from Baseline to Post-treatment

To evaluate within-person change from baseline to post-treatment, we estimated a multilevel model for each outcome. For example, the model for change in responses to the item "I feel happy" was specified like this:

Level 1:	$Happy_{ij} = \beta_{0j} + \beta_{1j}(Time)_{ij} + r_{ij}$
Level 2:	$\beta_{0j} = \gamma_{00} + \mu_{0j}$
	$\beta_{1j} = \gamma_{10} + \mu_{1j}$

Time was scored 0 (baseline) and 1 (post), so the model's intercept estimates people's scores at baseline, and the model's slope is the estimated change (in unstandardized, raw units) from baseline to post-treatment. Effects sizes were estimated using the proportional reduction in variance at the within-person level (Raudenbush and Bryk 2002), using the within-person variance (r_{ij}) for a null model as the baseline.

Item	Baseline value	Post-Tx change	Proportional variance reduction (%)	SE	р	Variance component
Cognition						
My thoughts are pleasant	3.66	.47	8.51	.17	.009	.82 (<i>p</i> < .001)
I have trouble concentrating	3.19	27	3.46	.12	.033	.36 (<i>p</i> < .001)
I am ruminating	3.23	32	10.36	.18	.090	.95 (<i>p</i> < .001)
Mood						
I feel happy	3.29	.62	14.41	.20	.004	1.29 (<i>p</i> < .001)
I feel confident	3.15	.88	11.02	.24	<.001	2.02 (<i>p</i> < .001)
I feel satisfied	3.17	.65	13.61	.20	.002	1.17 (<i>p</i> < .001)
I feel good about myself	3.37	.81	21.30	.21	<.001	1.44 (p < .001)
I feel like things are going well	3.50	.66	15.38	.20	.002	1.21 (<i>p</i> < .001)
I feel sad	3.47	63	13.36	.23	.008	1.49 (<i>p</i> < .001)
My emotions feel out of control	2.60	50	8.17	.17	.005	.91 (<i>p</i> < .001)
I feel anxious	3.50	31	6.65	.17	.084	.83 (<i>p</i> < .001)
I feel bad about myself	3.38	76	12.50	.17	<.001	.78 (<i>p</i> < .001)
I feel guilty	2.94	44	8.68	.16	.008	.70 (<i>p</i> < .001)
I feel irritable	3.40	43	6.61	.17	.013	.75 (<i>p</i> < .001)
I feel overwhelmed	3.78	30	6.61	.18	.108	.93 (<i>p</i> < .001)
Activities						
I like what I am doing	4.05	.37	4.38	.16	.021	.59 (<i>p</i> < .001)
I am successful in my current activity	4.67	.26	3.50	.12	.044	.36 (<i>p</i> < .001)
I would prefer to do something else	3.84	41	3.07	.16	.014	.58 (<i>p</i> < .001)
I am not doing what I should be doing	3.19	63	7.30	.17	<.001	.84 (<i>p</i> < .001)
Right now there is nothing I want to do	3.32	48	7.69	.19	.018	1.08 (<i>p</i> < .001)
Physical functioning						
I feel tired	4.45	50	8.83	.17	.006	.89 (<i>p</i> < .001)
I feel energetic	2.75	.43	13.04	.19	.027	1.09 (<i>p</i> < .001)
I don't feel well	3.26	64	14.27	.19	.002	1.24 (<i>p</i> < .001)
Social functioning						
Are you alone at this time?	61.06 %	-3.33 %	_	.17	.412	.58 (<i>p</i> < .001)
[When not alone] I feel close to this person (these people)	4.54	.16	5.23	.20	.428	.91 (<i>p</i> < .001)
[When not alone] Right now I would prefer to be alone	2.93	02	5.41	.20	.935	.90 (<i>p</i> < .001)
[When alone] I am alone right now because people do not want to be with me	1.89	37	5.33	.13	.008	.59 (<i>p</i> < .001)
[When alone] Right now I would prefer to be with other people	2.85	11	13.88	.26	.674	2.10 (<i>p</i> < .001)
Appraisal of situation						
My current situation is stressful	3.22	41	2.90	.12	.001	.29 (<i>p</i> < .001)
My current situation is positive	3 68	70	11.45	18	< 001	96(n < 001)

Table 3. Changes in daily life experiences with treatment

n = 55. See Table 2 for the instructions and response formats for the items. Baseline value is the estimated baseline value; Post-Tx change is the estimated change (slope) from baseline to post treatment. SE and p = standard error and p value for the change. Variance component = the variance component for the change slope, representing the magnitude of between-person heterogeneity in rates of change. Proportional variance reduction = the percentage decline in the within-person residual (r_{ij}) when time is included as a Level 1 predictor, compared to a null baseline model. "Are you alone at this time?" is a binary outcome, so the estimated probabilities of being alone (instead of the log-odds coefficients) are reported for ease of interpretation

Table 3 displays the effects. When evaluated in terms of statistical significance, the findings suggest that treatment had a broad effect across a range of measures of functioning. Items in most of the categories showed significant change in the expected direction. Apart from a handful of marginally significant items (including feeling momentarily anxious and overwhelmed), items assessing cognition, PA, NA, evaluations of current activities, and appraisals of the situation were statistically significant. Notably, only the social behavior category showed little change: apart from the item assessing whether people thought they were alone because others didn't want to be with them, none of the social behavior items changed significantly across time.

When evaluated in terms of effect sizes, many items showed substantial change, as shown in Fig. 2. (As a binary outcome, the item "Are you alone at this time?" lacks the residual variance needed to compute effect size). As the Proportional Variance Reduction (PVR) column in Table 3 shows, including time as a predictor reduced the unexplained variance by at least 10 % for 12 of the 29 items, and only 5 of 29 items had effects accounting for less than 5 % of the variance. These 5 items primarily involve ratings of current activities. Of note, the effects for the PA items were greater than for the NA items and among the largest effects overall. All 5 of the PA items had effects greater than 10 %, and the item with the single largest effect size—"I feel good about myself," at 23 %—reflects positive self-evaluation.



Figure 2. Effects of psychotherapy on daily experience in descending effect-size order

Moderators of Change by Condition

Our next models evaluated whether the rates of change were moderated by treatment condition. As an example of how the moderation models were specified, the model for whether happiness changed more in one condition used this specification:

Level 1: Happy_{ij} = $\beta_{0j} + \beta_{1j}$ (Time)_{ij} + r_{ij} Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01}$ (Treatment Condition)_j + μ_{0j} $\beta_{1j} = \gamma_{10} + \gamma_{11}$ (Treatment Condition)_j + μ_{1j}

Treatment condition was scored 0 for CBT and 1 for SST.

No statistically significant effects appeared for any of the items. There was one marginal effect for ratings of how close people felt to the people they were with ($\gamma = -.68$, SE = .38, p = .080), but this interaction was driven by a baseline difference.² Overall, this pattern suggests that the broad effects of psychotherapy on functioning were not specific to a particular therapy. Furthermore, this similarity in ESM markers of functioning parallels the findings for the weekly BDI-II and BAI assessments, which showed that the CBT and SST conditions had similar trajectories of change (Eddington et al. 2015).

Evaluating Affective Reactivity

Our final set of models examined changes in affective reactivity. We estimated two sets of models that examined whether time moderated the effect of being in a stressful situation or in a positive situation on emotions and thoughts. For example, being in a stressful situation should predict feeling overwhelmed, but the strength of the relationship should be weaker after therapy than before. The following equations illustrate such a model:

Level 1: Overwhelmed = $\beta_{0j} + \beta_{1j}$ (Time)_{ij} + β_{2j} (Stressful)_{ij} + β_{3j} (Interaction)_{ij} + r_{ij} Level 2: $\beta_{0j} = \gamma_{00} + \mu_{0j}$ $\beta_{1j} = \gamma_{10} + \mu_{1j}$ $\beta_{2j} = \gamma_{20} + \mu_{2j}$ $\beta_{3j} = \gamma_{30} + \mu_{3j}$

Stress was measured with the item "My current situation is stressful" and was centered at each person's own mean. An interaction term was formed by multiplying the stress scores with time, scored 0 for baseline and 1 for post-treatment. For outcomes, we tested all the items in the Cognition and Affect categories. A similar approach was used for the item "My current situation is positive."

For stressful situations, the models revealed statistically significant within-person interactions for most of the items. Specifically, time (baseline to post therapy) significantly moderated the effect of stress on all of the NA items and on the cognition items reflecting having trouble concentrating and ruminating (see Table 4). Notably, the effects were relatively specific to NA

² Specifically, at baseline, people who were assigned to CBT had lower values (M = 4.34) than people assigned to SST (M = 4.84), but the means were similar post-treatment (CBT M = 4.78, SST M = 4.60).

and experiences. The items associated with PA and thoughts were either marginal or not significant.

Outcome	Stressful situation		Positive situat	ion	
	b (SE)	р	b (SE)	р	
Нарру	.08 (.05)	.133	03 (.05)	.516	
Confident	.02 (.05)	.672	.02 (.06)	.789	
Satisfied	.02 (.05)	.656	05 (.06)	.369	
Good about myself	.08 (.05)	.140	.00 (.04)	.938	
Going well	.07 (.05)	.149	03 (.05)	.587	
Sad	15 (.07)	.034*	.13 (.07)	.049*	
Out of control	13 (.05)	.011*	.03 (.05)	.627	ļ
Anxious	13 (.05)	.011*	.07 (.07)	.313	
Bad about myself	14 (.03)	<.001*	.07 (.05)	.135	
Guilty	16 (.04)	<.001*	.10 (.05)	.041*	
Irritable	17 (.07)	.009*	.03 (.07)	.689	
Overwhelmed	13 (.05)	.012*	.06 (.07)	.407	
Thoughts are positive	.01 (.06)	.846	01 (.07)	.886	
Trouble concentrating	10 (.05)	.022*	.01 (.06)	.878	
Ruminating	10 (.04)	.014*	.05 (.05)	.240	

Table 4. Changes in the effects of stressful and positive situations on mood and cognitions from baseline to post-treatment

The coefficients represent within-person interactions between time (0 = baseline, 1 = post-treatment) and ratings of the current situation (within-person centered). b = unstandardized regression weight; SE = standard error; p = p value. The full item wording is shown in Table 1 * p < .05

All of the significant and marginal interactions indicated improved resilience. For example, Fig. 3 depicts the pattern for the item "I feel bad about myself right now," which is typical for that domain. For each outcome, stress predicted stronger negative experiences. After treatment, however, the slopes between stress and negative experiences became weaker. Treatment thus increased people's resilience in the face of stress: although they viewed the situation as stressful, their thoughts and emotions were less negatively affected.



Figure 3. The within-person interaction of time, stressful situations, and feeling bad about oneself

For positive situations, the models revealed very few statistically significant within-person interactions. As shown in Table 4, time moderated the effect of positive situations on only two negative items (feeling sad and guilty). Among the positive items, none were even marginally significant. The pattern for these results, illustrated in Fig. 4 for the "I feel sad" item, suggests that overall sadness and guilt decreased with treatment and became less affected by being in situations that were less positive.



Figure 4. The within-person interaction of time, positive situations, and feeling sad

Discussion

This study is the first to our knowledge to use ESM to assess changes in daily life functioning within the context of a randomized controlled psychotherapy trial for depression. After therapy, participants reported significant improvements in affect and cognitions and (to a lesser extent) social functioning and activities. Consistent with our predictions and with our results using traditional symptom measures (Eddington et al. 2015), there were no major differences between the two treatment conditions, so our discussion focuses on overall results.

Although our primary focus in this study was on changes in daily life emotional experiences and reactivity to events, the ESM protocol assessed daily experiences in a number of domains. Among the "top 10" ESM items showing the largest proportional improvements from baseline to post-treatment (see Fig. 2), seven reflected increases in PA, positive situational appraisals, and positive aspects of self-concept (i.e., feeling good about oneself) while only three reflected decreases in negative experiences. In general, participants showed significant improvements in mood in the expected directions (increases in PA; decreases in NA). Self-reported daily anxiety (as well as feeling overwhelmed), however, did not change significantly from pre- to post-treatment. Given that the treatments were focused on depression, the limited impact on anxiety may not be surprising. Analyses using traditional, weekly symptom measures did show improvements in anxiety in this sample, although the effects were smaller compared to those for weekly measures of depression (Eddington et al. 2015).

We were also able to examine within-person interactions between situational appraisals and momentary affect and cognition. It is notable that, in terms of proportional variance reduction,

the items characterizing the current situation showed the most stability during treatment. Only modest improvements in situational appraisals (e.g., stressfulness, and liking what one is doing) were observed. However, significant improvement in negative reactivity (affect and cognitions) to stress occurred, consistent with previous ESM results showing similar improvements with medications (Wichers et al. 2009). After treatment participants felt less anxious, sad, and irritable under stress, and they reported less rumination and lower levels of feeling bad about themselves. This improvement was more dramatic under high (versus low) stress conditions. This pattern is not surprising—therapy cannot necessarily reduce the frequency of negative events but can improve the ability to cope. Previous studies are mixed concerning the extent to which depressed individuals show heightened baseline stress reactivity compared to controls (Bylsma et al. 2011; Myin-Germeys et al. 2003; Nezlek and Gable 2001; South and Miller 2014). However, our results suggest that differences in negative cognitions and affect that are observed in higher versus lower stress situations are attenuated after therapy. Because our study did not include a nondepressed group, we cannot evaluate the extent to which that change reflects a normalization of response.

What about affective changes in the context of positive situations? Participants did show improvement in the extent to which they perceived their current situation as positive, however, reactivity to those positive situations did not show substantial change. There was some improvement in the extent to which people felt sad and guilty in positive situations, but there was no change in PA or cognition. To the extent that there may be a mood brightening effect in depression (Bylsma et al. 2011), this effect does not appear to be enhanced with therapy. However, the increase in perceptions that one's current situation is positive may suggest that, by post-treatment, patients have increased the range of their daily activities and have more opportunities to experience PA in daily life.

Regarding changes in the cognitive domain, we saw moderate improvements in thought content at post-treatment, characterized by more pleasant thoughts and less guilt; improvement in rumination was marginal. The difficulty in reducing rumination with CBT has been noted previously (Watkins et al. 2011), and SST does not specifically target rumination. Interventions that focus on rumination as a treatment target may be associated with greater improvements in that domain (Watkins et al. 2011). The item assessing difficulties concentrating showed significant change with treatment but, in effect size terms, relatively little improvement. This finding is consistent with research showing that nonspecific cognitive impairments in depressed patients tend to be persistent (Reppermund et al. 2009; Roca et al. 2015).

Compliance with ESM procedures was variable but comparable to studies with clinical and nonclinical samples (Johnson et al. 2009; Silvia et al. 2013). Our procedures were rigorous compared to paper-and-pencil methods in that our participants had a narrow response window (10 min), and we could accurately verify the timing of responses using the IVR system. At post-treatment, for those who completed treatment, the mean response rate dropped, but the distribution of response rates at post-treatment shows that over half of the participants completed at least two-thirds of the surveys.

The generalizability of these findings may be limited due both to the small sample size and the fact that ESM is demanding and may attract more motivated participants (Palmier-Claus et

al. 2011). Additionally, not all participants completed treatment, and therefore we have posttreatment ESM data on only a subset of the full sample. Our drop-out rate of 34 % is on the higher end of, but comparable to, rates reported in other depression treatment studies (Hans and Hiller 2013), suggesting that the ESM procedures did not necessarily adversely impact retention. Furthermore, of the 56 participants allocated to treatment, only one was excluded for completing only a single ESM survey and 49 (88 %) returned for a first treatment session. Finally, a limitation inherent in ESM studies is the need to reduce participant burden; as such, each phone survey had to be extremely brief. To cover a broad array of domains, we relied on single item assessments.

The use of momentary assessment tools allows for more fine-grained analyses of variability in daily functioning and contextual factors and are less subject to retrospective reporting biases that may influence responses to traditional symptom measures. Although this study yielded both momentary and summative data (i.e., BDI-II scores), the varying time intervals between measurements as well as missing data complicate analyses aimed at determining the extent to which those measures correspond, especially given the small sample size.

In the current study we found that brief, structured psychotherapy was associated with increases in PA and experiences, decreases in NA and experiences, and improvements in stress reactivity in daily life. We did not find differential effects in the two treatment condition, which may be due in part to the small sample size. However, this lack of differentiation may reflect a common mechanism of change, a notion that is consistent with transdiagnostic models of mood and anxiety disorders. For example, a recent transdiagnostic model of mood and anxiety disorders proposes that affective experiences (increased positive and decreased negative affect) and emotion regulation are central to these disorders, and effective treatments are those that enhance the ability to cope with NA (e.g., stress reactivity) and enhance affective experiences (increase PA and decrease NA) (Hofmann et al. 2012). Our pattern of results is certainly consistent with this notion.

As advances in the technology supporting experience sampling methods continues to improve, the potential of these methods for exploring even more sophisticated questions, such as those about psychotherapy process and mechanisms of action, will become increasingly apparent. In addition, future studies should also evaluate the potential therapeutic effects of intensive daily monitoring procedures, such as those used in the current study.

Acknowledgments. This research was supported by Award Number R21MH090414 from the National Institute of Mental Health (Trial Registration Number NCT02134678). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health. We thank Tamara E. Foxworth and Ariana Hoet for their assistance with data collection.

Conflict of Interest. Kari M. Eddington, Chris J. Burgin, Paul J. Silvia, Niloofar Fallah, Catherine Majestic and Thomas R. Kwapil declare that they have no conflict of interest.

Ethical Approval. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent. All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (national and institutional). Informed consent was obtained from all individual subjects participating in the study.

Animal Rights. No animal studies were carried out by the authors for this article.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.—text revision). Washington, DC: American Psychiatric Publishing, Inc. Retrieved from <u>http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1994-97698-000&site=ehost-live</u>

Barge-Schaapveld, D. Q. C. M., & Nicolson, N. A. (2002). Effects of antidepressant treatment on the quality of daily life: An experience sampling study. *Journal of Clinical Psychiatry*, *63*(6), 477–485.

Barge-Schaapveld, D. Q. C. M., Nicolson, N. A., van der Hoop, R. G., & DeVries, M. W. (1995). Changes in daily life experience associated with clinical improvement in depression. *Journal of Affective Disorders*, *34*(2), 139–154. doi: <u>10.1016/0165-0327(95)00012-</u><u>C</u>.

Barrantes-Vidal, N., Chun, C. A., Myin-Germeys, I., & Kwapil, T. R. (2013). Psychometric schizotypy predicts psychotic-like, paranoid, and negative symptoms in daily life. *Journal of Abnormal Psychology*, *122*(4), 1077–1087. doi: <u>10.1037/a0034793</u>.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, *56*(6), 893–897.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.

Ben-Zeev, D., & Young, M. A. (2010). Accuracy of hospitalized depressed patients' and healthy controls' retrospective symptom reports: An experience sampling study. *Journal of Nervous and Mental Disease, 198*(4), 280–285.

Ben-Zeev, D., Young, M. A., & Madsen, J. W. (2009). Retrospective recall of affect in clinically depressed individuals and controls. *Cognition and Emotion*, 23(5), 1021–1040.

Burgin, C. J., Silvia, P. J., Eddington, K. M., & Kwapil, T. R. (2012). Palm or cell? Comparing personal digital assistants and cell phones for experience sampling research. *Social Science Computer Review*, *31*(2), 244–251. doi: 10.1177/0894439312441577.

Bylsma, L. M., Taylor-Clift, A., & Rottenberg, J. (2011). Emotional reactivity to daily events in major and minor depression. *Journal of Abnormal Psychology*, *120*(1), 155–167. doi: 10.1037/a0021662.

Conner, T. S., Tennen, H., Fleeson, W., & Barrett, L. F. (2009). Experience sampling methods: A modern idiographic approach to personality research. *Social and Personality Psychology Compass, 3*(3), 292–313. doi: 10.1111/j.1751-9004.2009.00170.x.

Depp, C. A., Kim, D. H., de Dios, L. V., Wang, V., & Ceglowski, J. (2012). A pilot study of mood ratings captured by mobile phone versus paper-and-pencil mood charts in bipolar disorder. *Journal of Dual Diagnosis*, 8(4), 326–332. doi: <u>10.1080/15504263.2012.723318</u>.

Dozois, D. J. A., Dobson, K. S., & Ahnberg, J. L. (1998). A psychometric evaluation of the Beck Depression Inventory-II. *Psychological Assessment*, 10(2), 83–89.

Ebner-Priemer, U. W., & Trull, T. J. (2009). Ambulatory assessment: An innovative and promising approach for clinical psychology. *European Psychologist*, *14*(2), 109–119. doi: <u>10.1027/1016-9040.14.2.109</u>.

Eddington, K. M., Silvia, P. J., Foxworth, T. E., Hoet, A., & Kwapil, T. R. (2015). Motivational deficits differentially predict improvement in a randomized trial of self-system therapy for depression. *Journal of Consulting and Clinical Psychology*, *83*(3), 602–616. doi: <u>10.1037/a0039058</u>.

Fisher, C. D., & To, M. L. (2012). Using experience sampling methodology in organizational behavior. *Journal of Organizational Behavior*, *33*(7), 865–877. doi: <u>10.1002/job.1803</u>.

Forbes, E. E., Stepp, S. D., Dahl, R. E., Ryan, N. D., Whalen, D., Axelson, D. A., et al. (2012). Real-world affect and social context as predictors of treatment response in child and adolescent depression and anxiety: An ecological momentary assessment study. *Journal of Child and Adolescent Psychopharmacology*, 22(1), 37–47. doi: 10.1089/cap.2011.0085.

Fydrich, T., Dowdall, D., & Chambless, D. L. (1992). Reliability and validity of the Beck Anxiety Inventory. *Journal of Anxiety Disorders*, 6(1), 55–61. doi: <u>10.1016/0887-6185(92)90026-4</u>.

Geschwind, N., Peeters, F., Drukker, M., van Os, J., & Wichers, M. (2011). Mindfulness training increases momentary positive emotions and reward experience in adults vulnerable to depression: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *79*(5), 618–628. doi: 10.1037/a0024595.

Hans, E., & Hiller, W. (2013). Effectiveness of and dropout from outpatient cognitive behavioral therapy for adult unipolar depression: A meta-analysis of nonrandomized effectiveness studies. *Journal of Consulting and Clinical Psychology*, *81*(1), 75–88. doi: <u>10.1037/a0031080</u>.

Hartmann, J. A., Wichers, M., Menne-Lothmann, C., Kramer, I., Viechtbauer, W., Peeters, F., et al. (2015). Experience sampling-based personalized feedback and positive affect: A randomized controlled trial in depressed patients. *PLoS ONE*, *10*(6), e0128095. doi: <u>10.1371/journal.pone.0128095</u>.

Hofmann, S. G., Sawyer, A. T., Fang, A., & Asnaani, A. (2012). Emotion dysregulation model of mood and anxiety disorders. *Depression and Anxiety*, 29(5), 409–416. doi: <u>10.1002/da.21888</u>.

Johnson, E. I., Grondin, O., Barrault, M., Faytout, M., Helbig, S., Husky, M., et al. (2009). Computerized ambulatory monitoring in psychiatry: A multi-site collaborative study of acceptability, compliance, and reactivity. *International Journal of Methods in Psychiatric Research, 18*(1), 48–57. doi: 10.1002/mpr.276.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613. doi: <u>10.1046/j.1525-1497.2001.016009606.x</u>.

Moore, T. M., Seavey, A., Ritter, K., McNulty, J. K., Gordon, K. C., & Stuart, G. L. (2014). Ecological momentary assessment of the effects of craving and affect on risk for relapse during substance abuse treatment. *Psychology of Addictive Behaviors, 28*(2), 619–624. doi: <u>10.1037/a0034127</u>.

Myin-Germeys, I., Peeters, F., Havermans, R., Nicolson, N. A., deVries, M. W., Delespaul, P., et al. (2003). Emotional reactivity to daily life stress in psychosis and affective disorder: An experience sampling study. *Acta Psychiatrica Scandinavica*, *107*(2), 124–131. doi: <u>10.1034/j.1600-0447.2003.02025.x</u>.

Naylor, M. R., Keefe, F. J., Brigidi, B., Naud, S., & Helzer, J. E. (2008). Therapeutic interactive voice response for chronic pain reduction and relapse prevention. *Pain*, *134*(3), 335–345. doi: <u>10.1016/j.pain.2007.11.001</u>.

Nezlek, J. B., & Gable, S. L. (2001). Depression as a moderator of relationships between positive daily events and day-to-day psychological adjustment. *Personality and Social Psychology Bulletin, 27*(12), 1692–1704. doi: 10.1177/01461672012712012.

Palmier-Claus, J. E., Myin-Germeys, I., Barkus, E., Bentley, L., Udachina, A., Delespaul, P. A. E. G., et al. (2011). Experience sampling research in individuals with mental illness: Reflections and guidance. *Acta Psychiatrica Scandinavica*, *123*(1), 12–20.

Parkinson, B., Briner, R. B., Reynolds, S., & Totterdell, P. (1995). Time frames for mood: Relations between momentary and generalized ratings of affect. *Personality and Social Psychology Bulletin, 21*, 331–339.

Peeters, F., Berkhof, J., Rottenberg, J., & Nicolson, N. A. (2010). Ambulatory emotional reactivity to negative daily life events predicts remission from major depressive disorder. *Behaviour Research and Therapy*, *48*(8), 754–760. doi: <u>10.1016/j.brat.2010.04.008</u>.

Peeters, F., Nicolson, N. A., Berkhof, J., Delespaul, P., & deVries, M. (2003). Effects of daily events on mood states in major depressive disorder. *Journal of Abnormal Psychology*, *112*(2), 203–211.

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1(3), 385–401.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage. Retrieved

from https://books.google.com/books?hl=en&lr=&id=uyCV0CNGDLQC&oi=fnd&pg=PR17&d q=raudenbush+hierarchical+linear+models+2002&ots=qA4NTo04QH&sig=cRKsXiQfDOVHK nLr3MlR66Uf1a4 Reppermund, S., Ising, M., Lucae, S., & Zihl, J. (2009). Cognitive impairment in unipolar depression is persistent and non-specific: Further evidence for the final common pathway disorder hypothesis. *Psychological Medicine*, *39*(4), 603–614. doi: <u>10.1017/S003329170800411X</u>.

Roca, M., Monzón, S., Vives, M., López-Navarro, E., Garcia-Toro, M., Vicens, C., et al. (2015). Cognitive function after clinical remission in patients with melancholic and non-melancholic depression: A 6 month follow-up study. *Journal of Affective Disorders, 171*, 85–92. doi: <u>10.1016/j.jad.2014.09.018</u>.

Rose, G. L., Skelly, J. M., Badger, G. J., Naylor, M. R., & Helzer, J. E. (2012). Interactive voice response for relapse prevention following cognitive–behavioral therapy for alcohol use disorders: A pilot study. *Psychological Services*, *9*(2), 174–184. doi: 10.1037/a0027606.

Schrader, G., Davis, A., Stefanovic, S., & Christie, P. (1990). The recollection of affect. *Psychological Medicine*, 20(1), 105–109. doi: 10.1017/S0033291700013271.

Shacham, S. (1983). A shortened version of the Profile of Mood States. *Journal of Personality Assessment, 47*(3), 305–306.

Silvia, P. J., Kwapil, T. R., Eddington, K. M., & Brown, L. H. (2013). Missed beeps and missing data: Dispositional and situational predictors of nonresponse in experience sampling research. *Social Science Computer Review*, *31*(4), 471–481.

South, S. C., & Miller, M. L. (2014). Measuring momentary stress, affect, and cognition: Relationships with the internalizing and externalizing spectra. *Journal of Psychopathology and Behavioral Assessment*, *36*(1), 93–104. doi: <u>10.1007/s10862-013-9365-2</u>.

Strauman, T. J., Vieth, A. Z., Merrill, K. A., Kolden, G. G., Woods, T. E., Klein, M. H., et al. (2006). Self-system therapy as an intervention for self-regulatory dysfunction in depression: A randomized comparison with cognitive therapy. *Journal of Consulting and Clinical Psychology*, *74*(2), 367–376.

Tidey, J. W., Monti, P. M., Rohsenow, D. J., Gwaltney, C. J., Miranda, R. J., McGeary, J. E., et al. (2008). Moderators of naltrexone's effects on drinking, urge, and alcohol effects in non-treatment-seeking heavy drinkers in the natural environment. *Alcoholism, Clinical and Experimental Research*, *32*(1), 58–66. doi: <u>10.1111/j.1530-0277.2007.00545.x</u>.

van Hooff, M. L. M., Geurts, S. A. E., Kompier, M. A. J., & Taris, T. W. (2007). "How fatigued do you currently feel?" Convergent and discriminant validity of a single-item fatigue measure. *Journal of Occupational Health, 49*(3), 224–234.

Vieth, A. Z., Strauman, T. J., Kolden, G. G., Woods, T. E., Michels, J. L., & Klein, M. H. (2003). Self-system therapy (SST): A theory-based psychotherapy for depression. *Clinical Psychology: Science and Practice*, *10*(3), 245–268.

Watkins, E. R., Mullan, E., Wingrove, J., Rimes, K., Steiner, H., Bathurst, N., et al. (2011). Rumination-focused cognitive-behavioural therapy for residual depression: Phase II randomised controlled trial. *The British Journal of Psychiatry*, *199*(4), 317–322. doi: <u>10.1192/bjp.bp.110.090282</u>. 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. doi: 10.1037/0022-3514.54.6.1063.

Wichers, M. C., Barge-Schaapveld, D. Q. C. M., Nicolson, N. A., Peeters, F., de Vries, M., Mengelers, R., et al. (2009). Reduced stress-sensitivity or increased reward experience: The psychological mechanism of response to antidepressant medication. *Neuropsychopharmacology*, *34*(4), 923–931. doi: <u>10.1038/npp.2008.66</u>.