How Anxious are You Right Now? Using Ecological Momentary Assessment to Evaluate the Effects of Cognitive Bias Modification for Social Threat Interpretations

Katharine E Daniel

Department of Psychology

University of Virginia

Pre-Dissertation Paper

Readers

Bethany A. Teachman

Gerald L. Clore

Abstract

This study examines the effectiveness of a one-week period of online Cognitive Bias Modification for Interpretations (CBM-I) for socially anxious individuals. In addition to measuring intervention effectiveness through traditional trait measures, this study investigates whether associated state measures are sensitive to intervention effects in daily life. One-hundred and six participants scoring high on a measure of trait social anxiety completed two in-lab sessions separated by five weeks of ecological momentary assessment, with $n = 51$ participants randomly assigned to receive the online CBM-I intervention half-way through the five-week monitoring period. In addition to examining change on traditional trait questionnaire outcomes, state outcomes were assessed by comparing participants' responses to randomly-timed and nightly surveys delivered to their mobile phones prior to and following the online intervention. Results indicate that the CBM-I training was more effective than monitoring alone in reducing trait negative interpretation bias, indicating target engagement. However, this change was not reliably accompanied by changes on other cognitive processing style outcomes, at either trait or state levels. Further, while trait and state social anxiety symptoms and fear of negative evaluation improved, these changes were not unique to the CBM-I intervention group. This study demonstrates the challenges and opportunities associated with investigating intervention effects in daily life.

How Anxious are You Right Now? Using Ecological Momentary Assessment to Evaluate the Effects of Cognitive Bias Modification for Social Threat Interpretations

Social situations are often ambiguous. Did your coworker not wave back because she failed to see you, or because she was uninterested in talking to you? Individuals with social anxiety disorder (SAD) have been found to be more likely than non-anxious individuals to interpret ambiguous situations as socially threatening (e.g., Amir, Beard, & Bower, 2005). Moreover, negative interpretation bias is expected to be particularly influential in the development and maintenance of the disorder, given the ambiguity inherent in most social interactions (Kuckertz & Amir, 2014). In fact, cognitive models of SAD suggest that by reducing the tendency to interpret ambiguous situations as threatening, social anxiety symptoms will improve (Hofmann, 2000, 2007; Rapee & Heimberg, 1997). In line with these models, Cognitive Behavioral Therapy (CBT), an evidence-based treatment that targets distorted interpretations via cognitive restructuring, has been linked to reductions in negative interpretation biases (Williams et al., 2015; and see Franklin, Huppert, Langner, Leiberg, & Foa, 2005). Further, change in interpretation bias has been shown to predict or mediate symptom improvement in CBT (Goldin et al., 2012; Teachman, Marker, & Clerkin, 2010). This suggests that reduced interpretation biases might represent a core mechanism through which effective treatments in SAD function.

While CBT is empirically supported for the treatment of SAD (e.g., Hofmann, Asnaani, Vonk, Sawyer, & Fang, 2012), and appears to target interpretation bias in treatment responders (Franklin et al., 2005), it is not universally effective or accessible. In addition to treatment barriers that are common across many disorders (e.g., time, cost, accessibility; Beard, 2011), the fear of negative evaluation and social avoidance that characterizes SAD can make it especially difficult for socially anxious people to seek treatment (Alonso et al., 2018; Grant et al., 2005; Olfson et al., 2000). In fact, fewer than half of individuals with SAD receive treatment and people often suffer for more than ten years before seeking help (Anxiety and Depression

Association of America). Thus, it is vital to translate active treatment mechanisms, like negative interpretation bias, into intervention delivery approaches that increase access to care. This study therefore examines how an online cognitive bias modification program designed to reduce negative interpretation bias is associated with changes on SAD-related outcomes, both in general and in daily life.

Cognitive Bias Modification for Interpretations

Cognitive Bias Modification for Interpretations (CBM-I) represents an alternative, targeted approach to intervening on interpretation bias to threat that can be delivered online, by computer, tablet, or mobile phone. Although CBM-I can take different forms (see Schoth & Liossi, 2017 for a review), the current paper uses an ambiguous scenarios training approach (adapted from Mathews & Mackintosh, 2000) that presents people with a series of sentences that outline disorder-relevant, ambiguous scenarios (i.e., in SAD, CBM-I scenarios are related to ambiguity about social performance or negative evaluation from others). Individuals resolve the emotional meaning of each scenario by solving the last word, which is presented as a word fragment. To reduce a rigid, negative interpretation bias, CBM-I trains people to resolve *most* of the ambiguous situations positively. Thus, by employing a scalable technology to directly target this cognitive bias, CBM-I is positioned to potentially help reduce the treatment gap (Chisholm et al., 2016) by offering an alternative (or adjunct) to in-person care.

Many studies have investigated the efficacy of CBM-I at reducing negative interpretation biases and symptoms. Results from these investigations have been somewhat mixed, and meta-analyses show that the effect of CBM-I tends to be stronger on cognitive biases than on anxiety symptoms, and in laboratory-delivered CBM-I than in internet-delivered CBM-I (e.g., Hallion & Ruscio, 2011; Jones & Sharpe, 2017). Despite the larger effect sizes typically produced by CBM-I in controlled laboratory settings, there is evidence that internet-delivered CBM-I can have meaningful effects on interpretation biases and social anxiety symptoms. For example, Yang and colleagues (2017) showed that smartphone-delivered CBM-I (that reinforced positive interpretations 100% of the time) led to significantly lower threat interpretation biases than the alternative condition (that reinforced positive and negative interpretations equally). In an internet-delivered CBM-I pilot study for individuals diagnosed with SAD, 48% of the intervention completers no longer met diagnostic criteria at the end of the 8-week, unguided intervention. Results were maintained at 6-week follow up, though note that there was no control group (Brettschneider, Neumann, Berger, Renneberg, & Boettcher, 2015). In yet another study, two weeks of online CBM-I led to a greater reduction in social evaluative fear than a visuospatial training control condition (Hoppitt et al., 2014). Taken together, these results support the feasibility of CBM-I when disseminated through scalable and accessible technology, though internet-delivered CBM-I does not always promote symptom-level changes (e.g., De Voogd, Wiers, De Jong, Zwitser, & Salemink, 2018).

Effectiveness of CBM-I in Daily Life

Previous CBM outcome studies have tended to measure effectiveness using trait self-report questionnaires (e.g., State Trait Anxiety Inventory in CBM-I; Beard & Amir, 2010), and occasionally through psychophysiological measures (e.g., heart rate in CBM-I; Joormann, Waugh, & Gotlib, 2015) or neural responses (e.g., event-related potentials in attention bias modification; Nelson, Jackson, Amir, & Hajcak, 2017). By measuring interpretation bias and symptom changes at the trait level, as is standard in CBM-I studies, these attempts at understanding the effectiveness of CBM-I have largely ignored other cognitive processes that are associated with interpretation bias and mental health (e.g., cognitive flexibility and cognitive reappraisal; Everaert et al., 2017), as well as any potential state level changes following CBM-I (see Clarke et al., 2016 for one exception that looked at the effects of attention bias modification for insomnia on nightly sleep indices). Importantly, associations between results that are measured at the trait level are often weakly associated with those measured at the state level (e.g., affect, Brose, Voelkle, Lövdén, Lindenberger, & Schmiedek, 2015; emotion regulation, Kashdan & Nezlek, 2012). Thus, trait-level findings investigating the effectiveness of CBM-I may say little about how the intervention influences individuals in their daily lives.

Further, to consider only a narrow range of trait outcomes, without reference to associated state processes, is at odds with our clients' goals for treatment. People go to therapy because they want to make lifestyle changes. Socially anxious clients do not report that they overcame the huge cost and time barriers to accessing treatment in order to merely reduce their score on a questionnaire. Yet, current treatment outcome monitoring relies mainly on trait self-report questionnaires, which are subject to recall bias (Shiffman, Stone, & Hufford, 2008). Despite this known limitation, clinical scientists continue to ask clients to somehow average their remembered experiences across multiple contexts and situations, rather than directly assess clients' lived experiences as they unfold in real time. As a result, we know very little about how CBM-I affects the *daily lives* of socially anxious individuals.

There is some empirical basis to expect that interpretation biases influence in-the-moment processes. For example, Beard and Amir (2010) showed that trait negative interpretation bias mediated the relationship between trait social anxiety and state anxiety in response to an in-lab speech stressor task. However, one meta-analytic review did not find support for CBM-I in reducing vulnerability to a future stressor (Menne-Lothmann et al., 2014), which may suggest that CBM-I does not consistently buffer against in-the-moment stress responses to socially threatening situations.

Given the mixed state of the CBM-I literature, it is important to test whether or not CBM-I changes daily indicators of social anxiety. Further, Moore, Depp, Wetherell, and Lenze (2016) recently showed that outcomes that were measured through EMA were more sensitive to a mindfulness-based stress reduction intervention than were the associated in-lab questionnaires, emphasizing the need to extend beyond in-lab trait questionnaires to make claims about interventions. Thus, this study aims to: 1) test the effectiveness of CBM-I in reducing *trait* social anxiety symptoms and cognitive styles that suggest selective, rigid threat interpretations (this set of analyses mirrors many prior CBM-I studies), and 2) test the effects of CBM-I in daily life on associated *state* variables captured through EMA (these analyses will allow for a novel evaluation of CBM-I's impact).

Overview and Hypotheses

Individuals scoring high on a measure of trait social anxiety completed two in-lab sessions separated by five weeks of mobile phone EMA monitoring. Approximately half of the participants were randomized to receive six online CBM-I training sessions during the third week of the monitoring portion of the study (i.e., half way through the study protocol). All participants completed questionnaires during both in-lab sessions to assess for trait-level changes in cognitive processing styles and social anxiety, and all participants completed up to six randomly-timed EMA surveys and one End of Day survey per day to assess for associated state-level changes following the CBM-I intervention.

We consider the EMA-only group, which completed the EMA monitoring protocol without access to CBM-I, to be an active control group given evidence that tracking your emotions and experiences can itself be helpful (e.g., routine outcome monitoring; see Boswell, Kraus, Miller & Lambert, 2015, for review). Further, Truong and colleagues (2017) provided recent qualitative support for the potential effectiveness of EMA in promoting adaptive changes. Specifically, when asked what they thought would be the effect of completing a multi-week EMA study, individuals at high risk for polydrug use reported that they expected EMAs would increase their self-reflection on their emotions and daily activities, and that increased self-reflection could help catalyze behavior change by encouraging individuals to alter unhelpful behaviors and reinforce helpful behaviors (Truong et al., 2017).

Hypotheses were pre-registered (https://osf.io/eprwt/) and a list of all hypothesized outcomes are available in Table 1. All hypotheses (H) at the trait level are listed first, then associated state-level hypotheses are provided.

[INSERT TABLE 1]

Trait Hypotheses.

In line with the expected target engagement, individuals assigned to CBM-I (vs. EMA-only control) are hypothesized to show a greater reduction in trait negative interpretation bias from baseline to follow-up (H1a). Further, by reducing rigid, negative interpretation bias, we expect that individuals assigned to CBM-I (vs. EMA-only) will show a greater increase in trait cognitive flexibility (H2a) and trait cognitive reappraisal (H2b) from baseline to follow-up. At the symptom level, individuals assigned to CBM-I (vs. EMA-only) are expected to show a greater reduction on trait social anxiety from initial pre-screen to follow-up (H3a) and on trait fear of negative evaluation from baseline to follow-up (H4a).

Weekly Hypothesis.

Individuals assigned to CBM-I (vs. EMA-only) are expected to show a greater reduction in EMA-assessed, weekly fear of negative evaluation from Weeks 1 and 2 to Weeks 4 and 5, given that the CBM-I intervention is delivered in Week 3 of the study (H4b). Unlike the other EMA-delivered surveys, which asked participants to report either on their in-the-moment experiences or over the entire day, this weekly survey asked participants to report the degree to which they remembered fearing negative evaluation throughout the previous week.

Daily Hypotheses.

Cognitive training away from rigid interpretations may enhance an individual’s ability to engage in cognitive reappraisal (based on a negative association observed between trait interpretation bias and trait use of cognitive reappraisal; Everaert et al., 2017). Thus, we expect that individuals assigned to CBM-I (vs. EMA-only) will show a greater increase in their daily self-rated ability to use cognitive reappraisal from Weeks 1 and 2 to Weeks 4 and 5 (H2c). Further, individuals assigned to CBM-I (vs. EMA-only) are expected to report a greater reduction in self-reported daily social avoidance (H3b) and daily self-reported worry over other people’s thoughts of them (H4c) from Weeks 1 and 2 to Weeks 4 and 5.

In-the-Moment Hypotheses.

During instances of relatively high personal negative affect (selected because these are times when emotion regulation is more likely to be needed than during non-distressed times; Tamir, John, Srivastava, & Gross, 2007), individuals assigned to CBM-I (vs. EMA-only) are expected to show a greater increase in the likelihood to report using some form of emotion regulation (vs. no attempt to regulate their negative affect) in their daily life from Weeks 1 and 2 to Weeks 4 and 5 (H2d). During instances in which some form of emotion regulation strategy use was reported, we expect that individuals assigned to CBM-I will show a relative increase in the likelihood to use cognitive reappraisal, specifically, than participants in the EMA-only control. To be comprehensive, given that we assessed for momentary use of 19 emotion regulation strategies as part of the larger study and that CBM-I may have effects on multiple strategies, we include exploratory analyses for the remaining strategies in the online supplement, but only lay out hypotheses for cognitive reappraisal.

Further, effective interventions are expected to reduce the intensity of in-the-moment anxiety that is characteristic of SAD. Thus, individuals assigned to CBM-I (vs. EMA-only) are expected to show a greater reduction in self-reports of state anxiety from Weeks 1 and 2 to Weeks 4 and 5 (H3c). Follow-up analyses will look specifically at momentary anxiety when individuals report being in a social situation versus being alone, given that social situations are expected to evoke a greater deal of anxiety for socially anxious individuals, and so CBM-I effects may be more pronounced under these conditions.

Methods

Participants

N = 114 participants consented to participate in the study in exchange for course credit and/or payment. Participants were eligible to enroll in the current study if they scored at or above a 29 on the Social Interaction Anxiety Scale (SIAS, Mattick & Clarke, 1998), which ranges from 0 to 80, where higher scores indicate greater symptom severity. The cutoff score of 29 was determined *a priori* to ensure participants were experiencing moderate to severe social anxiety symptoms prior to beginning the study, and represents approximately 25% of a standard deviation below the average score observed in a sample diagnosed with social phobia (*M* = 34.6, *SD* = 16.4; Mattick & Clarke, 1998). Participants were recruited through advertisements sent to university email listservs for undergraduate and graduate students, through a psychology research participant pool, and through community flyers and online postings recruiting "socially anxious individuals aged 18-45 to participate in a 5-week smartphone monitoring study". After enrolling, participants were randomly assigned to either the CBM-I intervention group (*n* = 59) or to the EMA-only group (*n* = 55).

Eight participants in the CBM-I group did not initiate the first CBM-I training session (*n* = 3 dropped out of the study prior to the Week 3 intervention, *n* = 5 declined to initiate CBM-I but remained in the study), and were subsequently removed from analyses, leaving a final intent to treat sample of *N* = 106 (*n* = 51 in the CBM-I group and *n* = 55 in the EMA-only group).[1] All 51 participants who completed at least one CBM-I session were included in analyses, regardless of how many additional CBM-I sessions they completed. See Figure 1 for the CONSORT diagram.

[INSERT FIGURE 1]

There were no significant demographic differences between the CBM-I group (*n* = 51) and the EMA-only group (*n* = 55) on age ($t(104) = -.73$, $p = .47$), gender ($X^2(1) = 4.51e^{-31}$, $p = 1$), race ($X^2(4) = 4.68$, $p = .32$), initial SIAS pre-screen value ($t(100) = -.01$, $p = .92$), or affiliation with the authors' university (e.g., community member, undergraduate student, or graduate student; $X^2(2) = 3.42$, $p = .18$). However, the eight participants who did not initiate the CBM-I intervention were significantly younger than the CBM-I participants who did start the intervention (*M* = 19.38 vs. *M* = 20.67, $t(28) = -2.33$, $p < .05$) and there was a significant difference in the racial composition between the starters and the non-starters ($X^2(3) = 11.31$, $p < .05$). Namely,

[1] Analyses were originally run on all participants who returned to the follow-up lab session, regardless of whether or not they initiated the CBM-I intervention. However, in line with previous studies that defined intent to treat samples as those who initiated treatment (Ji, Meyer, & Teachman, under review), we re-ran all analyses on the sample described above. Notably, the pattern of results did not change.

no Asian participants (out of 10) failed to start the CBM-I intervention, whereas three (out of 5) African American/black participants did not start the CBM-I intervention. There were no differences on gender ($X^2(1) = .13$, $p = .72$), initial SIAS pre-screen value ($t(8)=.26$, $p = .80$), or university affiliation ($X^2(2) = 3.16$, $p = .21$) between the starters and non-starters. See Table 2 for detailed demographic information on the 106 participants that were included in analyses.

[INSERT TABLE 2]

Measures

Lab-based trait measures.

Trait negative interpretation bias. Negative interpretation bias was assessed using the negative external events subscale of the Brief Body Sensations Interpretations Questionnaire (BBSIQ; Clark et al., 1997). (Note, participants also completed the panic/physical events subscale as part of the larger study, but that subscale was not central to the current hypotheses and is not reported here.) Participants were presented with a list of 7 situations tied to social and every-day life events, and three potential explanations for each situation. For example, one situation reads, “A friend suggests that you change the way that you're doing a job in your own house. Why?” Two of the three potential explanations are benign with respect to the threat in how they interpret the situation (e.g., “They are trying to be helpful” and “They have done the job more often and know an easier way”), while the third explanation assigns a threatening interpretation (e.g., “They think you’re incompetent”). Participants are instructed to rank the likelihood of each statement on a 0 (“*Not at all likely*”) to 8 (“*Extremely likely*”) scale. A negative interpretation bias of external events is calculated by averaging the negative interpretation expectancy score for all seven of the external events scenarios, where higher scores indicate a greater bias towards negative interpretations of ambiguous situations. Internal consistency ranged from excellent to good at both assessment points (α = .91 at baseline and α = .83 at follow-up) in the present sample.

Trait cognitive flexibility. Cognitive flexibility was assessed using five items from the Cognitive Flexibility Inventory Alternatives Subscale (CFI-A; Dennis & Vander Wal, 2010), which assesses trait tendencies towards considering multiple options and thinking about situations from different perspectives. Items were rated on a 1 ("*Strongly Disagree*") to 7 ("*Strongly Agree*") scale, with higher scores indicating greater cognitive flexibility (vs. rigidity) when considering multiple alternatives during difficult situations. The five included items demonstrated above a .70 factor loading onto the "alternatives" subscale in a large prior sample (Dennis & Vander Wal, 2010). Internal consistency was good at both assessment points (α = .89 at baseline and α = .85 at follow-up) in the present sample.

Trait cognitive reappraisal. The general tendency to use cognitive reappraisal as an emotion regulation strategy was measured using the Reappraisal Subscale of the Emotion Regulation Questionnaire (ERQ-R; Gross & John, 2003). The subscale is composed of six items measured on a 1 ("*Strongly Disagree*") to 7 ("*Strongly Agree*") scale, with higher scores indicating a greater tendency to use cognitive reappraisal. Internal consistency was excellent at both assessment points (α = .91 at baseline and α = .90 at follow-up) in the present sample.

Trait social anxiety symptoms. Symptoms of social anxiety were assessed using the Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998) at two time-points: once prior to enrolling in the study and once again at the post-intervention lab session. Participants rated their agreement with 20 statements on a 0 ("*not at all characteristic of me"*) to 4 ("*extremely characteristic of me"*) scale. Prior to beginning the 5-week study, the average SIAS score across the full sample was 46.5 (*SD* = 10.2), which is nearly one standard deviation above the average SIAS score observed in a sample of individuals diagnosed with social anxiety disorder (M = 34.6, SD = 16.4; Mattick & Clarke, 1998). Internal consistency was excellent at both assessment points (α = .96 at pre-screen and α = .90 at follow-up) in the present sample.

Trait fear of negative evaluation. Fear of negative evaluation was assessed using the straightforwardly worded items from the original Brief Fear of Negative Evaluation Scale (BFNE;

Leary, 1983), given that the straightforward items (BFNE-S) demonstrated higher validity and reliability in a clinical sample of socially anxious individuals (Rodebaugh et al., 2011). Participants completed the BFNE-S at both in-lab sessions, separated by approximately five weeks. Participants rated their agreement with eight statements on a 1 ("*Not at all characteristic of me*") to 5 ("*Extremely characteristic of me*") scale, with higher scores indicating a greater fear of negative evaluation. Internal consistency was excellent at both assessment points (α = .94 at baseline and α = .93 at follow-up) in the present sample.

Ecological momentary assessment measures.

End of week measure.

Weekly fear of negative evaluation. To increase reliability of estimating change over time and to ensure that we had a recent assessment point if individuals dropped out, the BFNE-S was also administered during each of the five End of Week EMA surveys using a past-week modified version of the scale (e.g., "Over the past week, I was afraid others would not approve of me"). Internal consistency was excellent at each of the five weekly assessment points, with Cronbach's alpha ranging from .92 to .95.

End of day measures.

Daily self-rated ability to use cognitive reappraisal. Participants responded to the question, "I was ______ able to think about situations differently to change my thoughts/feelings when I wanted to" using a 6-point Likert scale ranging from 1 ("*not at all")* to 5 ("*Extremely"*). Participants could also report "I did not try to change my thoughts/feelings today."

Daily social avoidance. Participants rated their social avoidance using the single survey item, "I avoided my social interactions today" according to a 0 ("*Not at all*") to 10 ("*Very much*") sliding scale.

Daily social-evaluative concern. Participants completed the statement "I was _____ with what people might think of me today" with a sliding scale ranging from 0 ("*Very worried*") to 10 ("*Very comfortable*").

Randomly timed measures.

State self-reported use of emotion regulation. To measure likelihood to use some form of emotion regulation, participants reported which of 19 different emotion regulation strategies they had engaged in over the 30-minutes before each RT survey prompt using a check-all-that-apply list. Participants could select more than one strategy at a time, write-in additional strategies, or report they had not tried to change their thoughts and feelings over the previous 30 minutes. Participants are considered to have engaged in some form of emotion regulation (i.e., received a “0” vs. a “1”) if they reported using at least one of the 19 strategies and did not select “not trying to change my thoughts/feelings” in response to the question, “Over the 30 minutes before the survey prompt, I tried to change my thoughts and feelings through….” Specifically, participants were considered to have used cognitive reappraisal if they selected “Thinking of the situation differently” in response to the same question. A list of all 19 strategies that were assessed is included in the online supplemental materials.

State anxiety. Participants rated how calm relative to anxious they felt at each RT survey using the single item, “Right now, I am feeling…,” with anchors ranging from 0 (“*Very calm*”) to 10 (“*Very anxious*”). State anxiety *in the context of social situations* is defined as responses that occurred when the participant also indicated either being “around others but not interacting with them” or “interacting with others.”

Procedure

The university’s ethics review board approved the study. Participants consented to take part in a five-week smartphone monitoring study to examine thoughts and feelings throughout daily life, as well as to examine the effectiveness of a one-week period of online CBM-I for socially anxious individuals. The study was composed of two in-lab sessions separated by approximately five weeks, each lasting approximately one and a half hours. At baseline, participants completed various trait inventories, including measures of interpretation bias, cognitive reappraisal tendencies, and social anxiety symptoms.[2] All participants completed the same measures approximately five weeks later to assess for changes on trait outcomes over time (deviations to the protocol are detailed in Appendix A).

EMA protocol.

MetricWire, a smartphone app, was installed on all participants' personal devices at the end of the baseline lab session. During the five weeks between lab sessions, MetricWire was programmed to deliver six Random Time (RT) surveys throughout each day, randomly between each two-hour block between 9am-9pm (i.e., once between 9-11am, once between 11am-1pm, etc.). The survey was designed to take approximately two minutes to complete and participants were instructed to complete it as soon as possible upon receiving the notification. The app was programmed to remind participants to complete the survey after 30 minutes, and, if participants failed to respond to the survey within 45-minutes, the survey disappeared. Participants were instructed to answer the RT survey questions with reference to *when the survey first appeared on their phone*. While this instruction may introduce some minimal recall bias into the survey responses, we made this design decision to maximize the sampling of a broad range of situations that are part of daily life, even though some situations may make immediate survey responses difficult (e.g., when a survey is randomly prompted in the middle of a conversation). MetricWire was also programmed to deliver a three-minute End of Day (EOD) survey at 10pm every night, which instructed participants to reflect on their day, in general. EOD surveys remained active for two hours and automatically closed at midnight if unanswered. Finally, MetricWire was also programmed to deliver an End of Week (EOW) survey at 8pm at the end of each of the five weeks to measure fear of negative evaluation throughout each week. EOW surveys were available for twenty-four hours. The EMA schedule throughout Weeks 1, 2, 4, and 5 was identical for all participants.

[2] A full list of measures that were included in the larger study can be obtained by contacting the first author. No variables beyond those that are reported were analyzed for the current paper.

CBM-I protocol.

In addition to completing the daily phone surveys, the CBM-I intervention group was instructed to complete the online ambiguous scenarios training program (following Mathews and Mackintosh, 2000) during the third week of the five-week study. The CBM-I program involved six online sessions, each including 30 ambiguous scenarios and taking approximately 15 minutes to complete. Ninety percent of the scenarios from each training session were resolved positively. Participants were encouraged to complete one session each day for six days straight during Week 3 of the study, though participants could stop at any point or take more than six days to complete the training program.

Participant Payment

Participants received either $25 or 1.5 credit hours for participating in each laboratory session. At the end of the second laboratory session, participants were compensated for the five-week monitoring portion of the study based on the proportion of EMA surveys (and CBM-I intervention sessions, if applicable) that they completed. Payment for the monitoring portion of the study ranged from $10 to $80, with detailed payment information included in Appendix A.

Plan for Analyses

Data were analyzed using R version 3.4.3 (R core team, 2013). There was no significant difference in drop-out rate by condition ($X^2(1) = 5.55e^{-30}$, $p = 1$), and all participants in the CBM-I condition who completed at least one online session were included in analyses. On average, participants in the CBM-I condition completed 5.39 out of the 6 possible sessions ($SD = 1.39$, median = 6, ranging from 1 to 6), indicating that not all participants received the full dose of the brief intervention. Group demographic differences (CBM-I vs. EMA-only and CBM-I starters vs. non-starters) were examined using Welch two sample t-test for continuous variables (e.g., age) and chi-squared for categorical variables (e.g., gender). Cronbach's alpha was calculated for all trait outcome measures.

Linear mixed-effects models were run to test the effects of time and condition on each trait measure using the "lme4" package in R (Bates, Maechler, Bolker, & Walker, 2015) and p-values were obtained using the "lmerTest" package in R (Kuznetsova, Brockhoff, & Christensen, 2017). Time (baseline vs. follow-up), condition (CBM-I vs. EMA-only), and their interaction were entered as fixed effects, and participants were treated as random effects with a random intercept. The group-by-time interaction was the fixed effect of interest. The same procedure, with the addition of treating time as a random slope, was applied for all EMA data with continuous outcomes. Generalized logistic regression in the "lme4" package was used for binary outcomes that did not show zero-inflation (i.e., likelihood to use some form of emotion regulation strategy vs. no attempt at emotion regulation), whereas we used the "glmmTBM" package in R (Brooks et al., 2017) to fit specific emotion regulation strategy models using a zero-inflated distribution, as each strategy was reported infrequently. Analogously to the trait outcomes, "time" in all of the EMA state models was treated as a binary predictor (i.e., either the EMA response occurred in the first two weeks prior to the intervention or the EMA response occurred in the final two weeks following the intervention). This approach is consistent with methods described by Moore and colleagues (2016).

To assess for changes in the likelihood to use some form of emotion regulation during instances of personally high negative affect, all responses to the question "I was feeling…" on a 0 ("Very negative") to 10 ("Very positive") scale were standardized within person. Observations that corresponded to less than or equal to $1/4^{th}$ of a standard deviation below their personal mean (e.g., representing instances of personally high negative affect according to an *a priori* cut-off; https://osf.io/eprwt/) were retained from the first two weeks and from the last two weeks of EMA responses. To statistically control for the possibility that raw negative affect shifted over the course of the 5-week study within person, we entered in the mean standardized score from the observations retained in the first two weeks and the mean standardized score from the observations retained in the last two weeks for each person as a fixed effect control. Because participants each responded to a different number of surveys, and therefore mean estimates are based on differing numbers of observations, the number of observations that contributed to each mean value was also entered into the model as an additional fixed effect control (as mean estimates are increasingly robust with additional contributing observations). Further, to investigate changes in the likelihood to use each of the specific 19 emotion regulation strategies, given that some form of emotion regulation was reported, we removed all observations from the above analysis where no form of emotion regulation was reported. Control variables (i.e., average standardized negative affect and number of contributing observations) were recalculated to describe the observations that were included in these models.

Because properly interpretable R^2 statistics are not given by liner mixed-effects regressions (see Peugh, 2010), we report effect sizes using partially standardized beta estimates. Following Lorah (2018), we standardized all continuous outcome variables to achieve partially standardized beta estimates, where standardized beta estimates are usually smaller, more conservative estimates of *r* (see Ferguson, 2009). Given that our models are composed of binary predictor variables and it is not meaningful to standardize categorical variables, we could not achieve completely standardized beta estimates (Lorah, 2018). Further, in the absence of partially standardized beta estimates, beta estimates for models testing binary outcomes can be interpreted similarly to estimates of likelihood. We also report model-level marginal and conditional r-squared values, which we calculated according to the delta method in the “MuMIn” package in R (Bartoń, 2018). Marginal R^2 values represent the amount of variance in the outcome that is explained by only the fixed effects in the model, whereas conditional R^2 values represent the amount of variance that is explained by both the random and the fixed effects.

Results

All fixed and random effects statistics and conditional R^2 values for models are reported in Tables 3 (trait outcomes) and 4 (state outcomes).

Trait Outcomes.

The expected significant decline for participants in the CBM-I condition, relative to the EMA-only condition, was observed on negative interpretation bias on the BBSIQ, indicating intervention-specific target engagement. However, there were no significant time, condition, or time-by-condition interaction effects on trait cognitive flexibility on the CFI-A or trait cognitive reappraisal on the ERQ-R. At the symptom level, both conditions showed comparable significant declines on trait social anxiety symptoms on the SIAS and trait fear of negative evaluation on the BFNE-S.

[INSERT TABLE 3]

Weekly Outcome.

When comparing weekly ratings of fear of negative evaluation from the first two weeks to the last two weeks on the past-week modified BFNE-S, there was no significant main effect for time (though the effect was approaching significance in the same direction as was observed in trait fear of negative evaluation). Effects did not differ between the CBM-I and the EMA-only groups.

Daily Outcomes.

According to EOD survey responses, on days when participants reported using cognitive reappraisal, there was a significant time-by-condition interaction effect on daily self-reported ability to change one's thoughts and feelings using cognitive reappraisal. Specifically, during the first two weeks, participants in the CBM-I group (vs. EMA-only group) reported significantly lower perceived ability to use cognitive reappraisal. After the intervention, during the last two weeks, there was no group difference on perceived daily ability to use cognitive reappraisal. Moreover, the likelihood for participants to report having used cognitive reappraisal throughout the day on EOD surveys did not change for either group over time or as a function of the CBM-I intervention. Taken together, these results show that although neither group was *more likely* to use cognitive reappraisal leading up to or following the intervention period, participants in the

CBM-I group *rated* their reappraisal attempts *as more effective* in the last two weeks compared to the first two weeks, which allowed their ratings to "catch up" to the ability level endorsed by individuals in the EMA-only group.

In terms of social anxiety symptoms, there were no significant time, condition, or time-by-condition interaction effects on daily reports of social avoidance, indicating that participants' self-reported, daily degree of social avoidance did not change over time or as a function of the CBM-I intervention. However, both conditions showed comparable declines on reports of daily social-evaluative concern.

In-the-Moment Outcomes.

When comparing the likelihood to report using some form of emotion regulation (regardless of the specific strategy used) to the likelihood to report no attempt at emotion regulation, during instances of personally high negative affect, there was no significant time-by-condition interaction or main effect for time. However, there was a significant main effect for condition, such that individuals in the CBM-I group were more likely to report not trying to regulate their emotions across all instances of high negative affect than were individuals in the EMA-only group. With specific regard to the use of cognitive reappraisal during instances of personally high negative affect, within all observations in which some form of emotion regulation strategy was reported, we observed no significant time, condition, or time-by-condition interaction effects on state likelihood to report using cognitive reappraisal. Fixed effect statistics for cognitive reappraisal and the remaining 18 emotion regulation strategies are included in Table S1 of the online supplement.

Although not unique to the CBM-I intervention, both conditions showed comparable significant declines on state anxiety in general, and specifically during social situations.

[INSERT TABLE 4]

Discussion

This study is the first investigation to our knowledge of CBM-I on changes in both more traditional trait cognitive styles and symptom measures, and in state outcomes in the daily lives of socially anxious individuals. In terms of cognitive style outcomes, with the exception of trait negative interpretation bias, which showed the expected greater decline for participants in the CBM-I condition relative to the EMA-only control condition, all *trait* cognitive style outcomes remained stable throughout the 5-week study, regardless of study condition. All *state* cognitive style outcomes also remained stable over time for both groups, except for daily ability to use cognitive reappraisal, which showed that the individuals in the CBM-I condition “caught up” to the EMA-only group over time after rating themselves as less able to use cognitive reappraisal in the first two weeks.

Meanwhile, in terms of symptom outcomes, both conditions showed comparable significant declines on trait social anxiety symptoms and trait fear of negative evaluation, and on their associated state outcomes (state anxiety and daily social-evaluative concern, respectively). We observed no significant changes over time on weekly fear of negative evaluation (though the effect was a non-significant trend in the same direction as trait fear of negative evaluation and daily social-evaluative concern), nor daily reports of social avoidance.

Minimal Condition Differences

Consistent with the goal of the intervention, we observed CBM-I-specific declines on trait negative interpretation bias (the measure of target engagement). Despite evidence for target engagement, the observed significant improvements on trait and state social anxiety and fear of negative evaluation outcomes were not unique to the CBM-I group. Notably, previous CBM-I intervention studies have also found comparable significant improvements for active CBM-I training and various comparison conditions on anxiety symptoms (Brettschneider et al., 2015; De Voogd et al., 2018; Namaky et al., in preperation), making it unclear as to whether these improvements reflect the natural course of anxiety, regression to the mean, expectancy effects, or an active effect of the comparison condition(s). Similarly, in the current study, it is difficult to determine what is driving the observed improvements across both conditions. It may be that the improvements across both trait and state symptom outcomes resulted from increased self-reflection brought about by the EMA monitoring protocol (see Boswell et al., 2015; Truong et al., 2017), but without a no-monitoring comparison group it is not possible to know (Barnett, van der Pols, & Dobson, 2005).

With the exception of negative interpretation bias and daily ability to use cognitive reappraisal, the cognitive style outcomes remained stable throughout the 5-week study, regardless of study condition. These additional outcomes are not typically included in CBM-I intervention studies, at either the trait or state level, despite their plausible relevance (e.g., negative interpretation bias is negatively associated with trait cognitive reappraisal; Joormann et al., 2015; and CBM-I aims to reduce rigidly negative interpretations of ambiguous social situations). Notably, the cognitive style outcome measures that were used in the current study were not specific to ambiguous social situations. Instead, they considered a broad range of social and non-social distressing situations. Thus, it is not clear whether a larger dose of the CBM-I intervention could bring about changes on these outcomes in future studies, whether an alternate measure that more closely matches the training domain would show effects, or whether reducing negative interpretation bias via CBM-I may just not increase flexibility in cognitive styles.

Intervention Features.

The minimal unique CBM-I effects may be due to features of the current study's intervention protocol. First, this study investigated a relatively low dose of the CBM-I intervention, with each of the possible six training sessions lasting 10 to 15 minutes and including fewer training scenarios than other procedures (e.g., 30 scenarios six times in the current study vs. 40 scenarios 8 times in Ji et al., under review). Notably, one meta-analysis showed that the effect of CBM-I on symptom improvement was stronger with increasing number of sessions (Menne-Lothmann et al., 2014). Thus, it is possible that the modest number of sessions, coupled with the shorter duration of each session, did not deliver a strong enough dose of the intervention to bring about the expected, intervention-specific changes.

Second, the current study was designed to have CBM-I participants complete one session each day for 6 days straight (although deviations from this protocol did occur and not all participants assigned to CBM-I completed all training sessions). Many studies that have shown uniquely positive effects of online CBM-I have been structured such that participants completed longer sessions that were spaced further apart (e.g., twice a week for 8 weeks, Brettschneider et al., 2015; five times over two weeks, Hoppitt et al., 2014; minumum of 48 hours between sessions for up to 8 sessions, Ji et al., under review, etc.), perhaps allowing for greater skill scaffolding and spaced practicing (although one meta-analysis that combined attention and interpretation training paradigms across symptom domains in children found no spacing effects on symptom change; Cristea, Mogoase, David & Cuijpers, 2015). Further, though instructed to complete the online sessions during the third week of the study, participants could continue to complete online sessions in the fourth and (in rare circumstances) fifth weeks of the study, so not all EMA responses contributing to the “post” time period occurred following the full dose of the intervention. As a result, these analyses may underrepresent the training program’s impact on daily life.

Finally, despite high compliance in comparison to other online intervention studies (e.g., 80.4% of participants completed all six CBM-I sessions in the current study vs. fewer than 10% of completers in some Massive Open Online Course [MOOC] interventions; Gütl, Rizzardini, Chang, & Morales, 2014), it is possible that the current study’s participants were not motivated to change their anxious thinking. Notably, participants in the current study did not enroll for the purpose of gaining treatment (rather, participants were told that the study examines how people think, feel, and act as they go about their daily lives, and they would be testing a computer program designed to help people think in new ways about situations that may make them anxious) and treatment-seeking has been associated with readiness to change (Krampe et al., 2017). Thus, this non-treatment-seeking sample (40% of which enrolled for class credit) may have had a lower desire to incorporate the training into their daily lives compared to previous treatment-seeking samples (e.g., Ji et al., under review).

Limitations

The results of this study should be interpreted in light of several limitations. Response rates for the EMA portion of the study were somewhat low (63.8% of all surveys were answered across the first and last two weeks of the study), though the current study's response rate is quite similar to the observed mean across 45 mobile phone EMA studies (M = 69.6%, SD = 22.8%, where an additional 65 studies did not report compliance; Berkel, Ferreira, & Kostakos, 2017). Further, total number of survey responses did not differ between study groups ($M_{EMA\text{-}only}$ = 128.73 and $M_{CBM\text{-}I}$= 127.94, t(98.340) = .08, p = .93), although the number of responses were lower in the final two weeks of the study than they were in the first two weeks for both groups (B = -14.69, p < .001). In addition to technical glitches that may have contributed to some missed surveys and surveys firing to participants' phones and subsequently timing out before participants could answer them (e.g., while they were in a meeting), the low response rates may also be due in part to participant burden. To increase compliance, future studies may want to reduce the number of randomly timed surveys they deliver each day or further increase the incentive structure. Especially for long EMA protocols, researchers may also want to more heavily incentivize survey completion in the latter vs. beginning part of the study. That said, multi-level models are robust to unbalanced designs (Maas & Hox, 2005) and numerous effects for time were observed, so it is unlikely that missing surveys can fully account for the null intervention effects.

Additionally, although this study assessed a far wider array of outcomes than is typical for CBM-I studies, there were some notable gaps and weaknesses in the assessments. In particular, the current study did not include a measure of state negative interpretation bias, which would allow researchers to investigate whether participants change their thinking about ambiguous social situations in daily life as a function of the intervention. Given that CBM-I aims to promote symptom change through this mechanism, this would be a valuable addition, but it is not obvious how to measure this process in a low-burden way during daily life. Further, it is notable that daily social avoidance remained stable despite all other trait and state social anxiety symptom measures improving over time for all participants. It is possible that an end-of-day self-report measure of social avoidance does not adequately capture the actual social avoidance behaviors that are core to social anxiety (e.g., entering into fewer public spaces, avoiding eye contact, etc.). Researchers may want to consider objectively measuring avoidance behaviors using passive sensing smartphone technology (e.g., GPS to track the frequency and types of locations visited over time).

Finally, though enrollment was open to community members, participants in this study were mostly undergraduate and graduate students. Also, participants had to be high in symptoms but did not have to be diagnosed with social anxiety disorder to be eligible for the current intervention study. Thus, it will be important to test whether these findings replicate in non-university affiliated samples and with clinical populations.

Conclusion

This study is the first investigation (to our knowledge) into the effects of CBM-I on state outcomes in the daily lives of socially anxious individuals. Results showed the expected CBM-I-specific decline on trait negative interpretation bias in a predominantly student sample scoring high on a measure of trait social anxiety. Despite target engagement, trait and state improvements on cognitive style and social anxiety symptom constructs, when observed, were not unique to the CBM-I intervention, relative to the EMA-only control condition. Given the importance of evaluating how an intervention affects *the daily lives* and relationships of the individuals it intends to benefit, future studies should continue to leverage EMA-methodology to explore intervention effects in the real world. This study demonstrates the challenges and opportunities that are associated with investigating intervention effects in daily life.

References

Alonso, J., Liu, Z., Evans-Lacko, S., Sadikova, E., Sampson, N., Chatterji, S., … Thornicroft, G. (2018). Treatment gap for anxiety disorders is global: Results of the World Mental Health Surveys in 21 countries. *Depression and Anxiety*. https://doi.org/10.1002/da.22711

Amir, N., Beard, C., & Bower, E. (2005). Interpretation bias and social anxiety. *Cognitive Therapy and Research, 29*(4), 433-443. https:/doi.org/10.1007/s10608-005-2834-5

Barnett, A.G., van der Pols, J.C., & Dobson, A.J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology, 34*(1), 215-20 https://doi.org/10.1093/ije/dyh299

Bartoń, K. (2018). MuMIn: Multi-model inference. R package. 1.43.6 See https://cran.r-project.org/web.packages/MuMIn/index.html

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi.org/10.18637/jss.v067.io1

Beard, C. (2011). Cognitive bias modification for anxiety: Current evidence and future directions. *Expert Rev Neurother. 11*(2), 299-311. https://doi.org/10.1586/ern.10.194

Beard, C., & Amir, N. (2010). Negative interpretation bias mediates the effect of social anxiety on state anxiety. *Cognitive Therapy and Research*. https://doi.org/10.1007/s10608-009-9258-6

Beard, C., Weisberg, R. B., & Primack, J. (2012). Socially anxious primary care patients attitudes toward cognitive bias modification (CBM): A qualitative study. *Behavioural and Cognitive Psychotherapy*, *40*(5), 618–633. https://doi.org/10.1017/S1352465811000671

Berkel, N. Van, Ferreira, D., & Kostakos, V. (2017). The Experience Sampling Method on Mobile Devices. *ACM Computing Surveys*. https://doi.org/10.1145/3123988

Brettschneider, M., Neumann, P., Berger, T., Renneberg, B., & Boettcher, J. (2015). Internet-based interpretation bias modification for social anxiety: A pilot study. *Journal of Behavior*

Therapy and Experimental Psychiatry. https://doi.org/10.1016/j.jbtep.2015.04.008

Boswell, J.F., Kraus, D.R., Miller, S.D., & Lambert, M.J. (20150. Implementing routine outcome monitoring in clinical practice: benefits, challenges, and solutions. *Psychother Res, 25*(1), 6-19. https://doi.org/10.1080/10503307.2013.817696

Brooks, M.E., Kristensen, K., van Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Maechler, M., & Bolker, B.M. (2017). glmmTBM balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal, 9*(2), 378-400.

Brose, A., Voelkle, M.C., Lövdén, M., Lindenberger, U., & Schmiedek, F. (2015). Differences in the between-person and within-person structures of affect are a matter of degree. *European Journal of Personality, 29*(1), 55-71. https://doi.org/10.1002/per.1961

Chisholm, D., Sweeney, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016). Scaling-up treatment of depression and anxiety: A global return on investment analysis. *The Lancet Psychiatry, 3*(5), 415-424. https://doi.10.1016/S2215-0366(16)30024-4

Clark, D.M., Salkovskis, P.M., Pst, L.G., Breitholtz, E., Koehler, K.A., Westling, B.E., & Gelder, M. (1997). Misinterpretation of body sensations in panic disorder. *Journal of Consulting and Clinical Psychology, 65*, 203-2013

Clarke, P. J., Bedford, K., Notebaert, L., Bucks, R. S., Rudaizky, D., Milkins, B. C., & MacLeod, C. (2016). Assessing the therapeutic potential of targeted attentional bias modification for insomnia using smartphone delivery. *Psychother Psychosom*, *85*(3), 187-189. https://doi.org/10.1159/000442025

Cristea, I. A., Kok, R. N., & Cuijpers, P. (2015). Efficacy of cognitive bias modification interventions in anxiety and depression: Meta-analysis. *British Journal of Psychiatry*. https://doi.org/10.1192/bjp.bp.114.146761

Cristae, I.A., Mogoase, C., David, D., & Cuijpers, P. (2015). Practitioner review: Cognitive bias modification for mental health problems in children and adolescents: A meta-analysis. The *Journal of Child Psychology and Psychiatry, 56*(7), 723-734. https://doi.org/10.1111/jcpp.12383

Dennis, J.P., & Vander Wal, J.S. (2010). The cognitive flexibility inventory: Instrument development and estimates of reliability and validity. C*ognitive Therapy and Research, 34*(3), 241-253. https://doi.org/10.1007/s10608-009-9276-4

De Voogd, L., Wiers, R. W., De Jong, P. J., Zwitser, R. J., & Salemink, E. (2018). A randomized controlled trial of multi-session online interpretation bias modification training: Short- and long-term effects on anxiety and depression in unselected adolescents. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0194274

Everaert, J., Grahek, I., Duyck, W., Buelens, J., Van den Bergh, N., & Koster, E. H. W. (2017). Mapping the interplay among cognitive biases, emotion regulation, and depressive symptoms. *Cognition and Emotion*. https://doi.org/10.1080/02699931.2016.1144561

Facts & Statistics. Retrieved from https://adaa.ord/about-adaa/press-room/facts-statistics

Ferguson, C. J. (2009). An Effect Size Primer: A Guide for Clinicians and Researchers. *Professional Psychology: Research and Practice*, *40*(5), 532–538. https://doi.org/10.1037/a0015808

Franklin, M. E., Huppert, J., Langner, R., Leiberg, S., & Foa, E. B. (2005). Interpretation bias: A comparison of treated social phobics, untreated social phobics, and controls. *Cognitive Therapy and Research*. https://doi.org/10.1007/s10608-005-2412-8

Goldin, P. R., Ziv, M., Jazaieri, H., Werner, K., Kraemer, H., Heimberg, R. G., & Gross, J. J. (2012). Cognitive reappraisal self-efficacy mediates the effects of individual cognitive-behavioral therapy for social anxiety disorder. *Journal of Consulting and Clinical Psychology*. https://doi.org/10.1037/a0028555

Gross, J. J., & John, O. P. (2003). Individual Differences in Two Emotion Regulation Processes: Implications for Affect, Relationships, and Well-Being. *Journal of Personality and Social*

Psychology. https://doi.org/10.1037/0022-3514.85.2.348

Gütl, C., Rizzardini, R.H., Chang, V., & Morales, M. (2014). Attrition in MOOC: Lessions learned from drop-out students. In: Uden, L., Sinclair, J., Tao, Y.H., Liberona, D. (eds) Learning Technology for Education in Cloud. MOOD and Big Data. LTEC 2014. Communications in Computer and Information Science, vol 446. Springer, Cham

Hallion, L. S., & Ruscio, A. M. (2011). A Meta-Analysis of the Effect of Cognitive Bias Modification on Anxiety and Depression. *Psychological Bulletin*. https://doi.org/10.1037/a0024355

Hofmann, S. G. (2000). Treatment of social phobia: Potential mediators and moderators. *Clinical Psychology: Science and Practice*, *7*, 3–16.

Hofmann, S. G. (2007). *Cognitive Factors that Maintain Social Anxiety Disorder: a Comprehensive Model and its Treatment Implications*.

Hofmann, S. G., Asnaani, A., Vonk, I. J. J., Sawyer, A. T., & Fang, A. (2012). The Efficacy of Cognitive Behavioral Therapy: A Review of Meta-analyses NIH Public Access. *Cognit Ther Res*, *36*(5), 427–440. https://doi.org/10.1007/s10608-012-9476-1

Hoppitt, L., Illingworth, J. L., MacLeod, C., Hampshire, A., Dunn, B. D., & Mackintosh, B. (2014). Modifying social anxiety related to a real-life stressor using online Cognitive Bias Modification for interpretation. *Behaviour Research and Therapy*. https://doi.org/10.1016/j.brat.2013.10.008

Jones, E. B., & Sharpe, L. (2017, December 1). Cognitive bias modification: A review of meta-analyses. *Journal of Affective Disorders*. Elsevier B.V. https://doi.org/10.1016/j.jad.2017.07.034

Joormann, J., Waugh, C. E., & Gotlib, I. H. (2015). Cognitive bias modification for interpretation in major depression: Effects on memory and stress reactivity. *Clinical Psychological Science*. https://doi.org/10.1177/2167702614560748

Kashdan, T.B. & Nezlek, J.B. (2012). Whether, when, and how is spirituality related to well-being? Moving beyond single occasion questionnaires to understanding daily process. *Personality and Social Psychology Bulletin, 38*, 1526-1538. https://doi.org/10.1177/0146167212454549

Krampe, H., Salz, A. L., Kerper, L. F., Krannich, A., Schnell, T., Wernecke, K. D., & Spies, C. D. (2017). Readiness to change and therapy outcomes of an innovative psychotherapy program for surgical patients: Results from a randomized controlled trial. *BMC Psychiatry*, *17*(1). https://doi.org/10.1186/s12888-017-1579-5

Leary, M. (1983). A brief version of the fear of negative evaluation scale. *Personality and Social Psychology Bulletin, 9*(3), 371-375, https://doi.org/10.1177/0146167283093007

Lorah, J. (2018). Effect size measures for multilevel models: definition, interpretation, and TIMSS example. *Large-Scale Assessments in Education*, *6*(1). https://doi.org/10.1186/s40536-018-0061-2

Ji, J.L., Meyer, M.J., & Teachman, B. (Under review). Facilitating episodic simulation in anxiety: Role of sensory scaffolding and scenario modality.

Maas, C. J. M., & Hox, J. J. (2005). Sufficient Sample Sizes for Multilevel Modeling. *Methodology*, *1*(3), 86–92. https://doi.org/10.1027/1614-1881.1.3.86

Mathews, A., & Mackintosh, B. (2000). Induced Emotional Interpretation Bias and Anxiety. *Journal of Abnormal Psychology*, *109*(4), 615. https://doi.org/10.1037//002I-843X.109.4.602

Mattick, R.P., & Clarke, J.C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behav Res Ther, 36*(4), 455-70

Menne-Lothmann, C., Viechtbauer, W., Höhn, P., Kasanova, Z., Haller, S. P., Drukker, M., … Lau, J. Y. F. (2014). How to boost positive interpretations? A meta-analysis of the effectiveness of cognitive bias modification for interpretation. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0100925

Moore, R. C., Depp, C. A., Wetherell, J. L., & Lenze, E. J. (2016). Ecological momentary assessment versus standard assessment instruments for measuring mindfulness, depressed mood, and anxiety among older adults. *Journal of Psychiatric Research*. https://doi.org/10.1016/j.jpsychires.2016.01.011

Nelson, B.D., Jackson, F., Amir, N., Hajcak, G. (2017). Attention bias modification reduces neural correlates of response monitoring. *Biol Psychol, 129*, 103-110. https://doi.org/10.1016/j.biopsycho.2017.08.059

Olfson, M., Shea. S., Feder, A., Fuentes, M., Nomura, Y., Gameroff, M., & Weissman, M.M. (2000). Prevalence of anxiety, depression, and substance use disorders in an urban general medicine practice. *Arch Fam Med, 9*(9), 876-83.

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*. https://doi.org/10.1016/j.jsp.2009.09.002

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy*. https://doi.org/10.1016/S0005-7967(97)00022-3

Rodebaugh, T. L., Heimberg, R. G., Brown, P. J., Fernandez, K. C., Blanco, C., Schneier, F. R., & Liebowitz, M. R. (2011). More reasons to be straightforward: Findings and norms for two scales relevant to social anxiety. *Journal of Anxiety Disorders*. https://doi.org/10.1016/j.janxdis.2011.02.002

Salemink, E., Kindt, M., Rienties, H., & Van Den Hout, M. (2014). Internet-based Cognitive Bias Modification of Interpretations in patients with anxiety disorders: A randomised controlled trial. *Journal of Behavior Therapy and Experimental Psychiatry*. https://doi.org/10.1016/j.jbtep.2013.10.005

Schoth, D. E., & Liossi, C. (2017). A systematic review of experimental paradigms for exploring biased interpretation of ambiguous information with emotional and neutral associations. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2017.00171

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological Momentary Assessment. *Annual Review of Clinical Psychology*. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Tamir, M., John, O. P., Srivastava, S., & Gross, J. J. (2007). Implicit Theories of Emotion: Affective and Social Outcomes Across a Major Life Transition. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.92.4.731

Teachman, B. A., Marker, C. D., & Clerkin, E. M. (2010). Catastrophic misinterpretations as a predictor of symptom change during treatment for panic disorder. *Journal of Consulting and Clinical Psychology*, *78*(6), 964–973. https://doi.org/10.1037/a0021067

Truong, Q., Roth, A. M., Simmons, J., Garfein, R. S., Goldshear, J. L., Felsher, M., & Reed, M. (2017). Potential benefits of using ecological momentary assessment to study high-risk polydrug use. *MHealth*. https://doi.org/10.21037/mhealth.2017.10.01

Williams, A. D., O'Moore, K., Blackwell, S. E., Smith, J., Holmes, E. A., & Andrews, G. (2015). Positive imagery cognitive bias modification (CBM) and internet-based cognitive behavioral therapy (iCBT): A randomized controlled trial. *Journal of Affective Disorders*. https://doi.org/10.1016/j.jad.2015.02.026

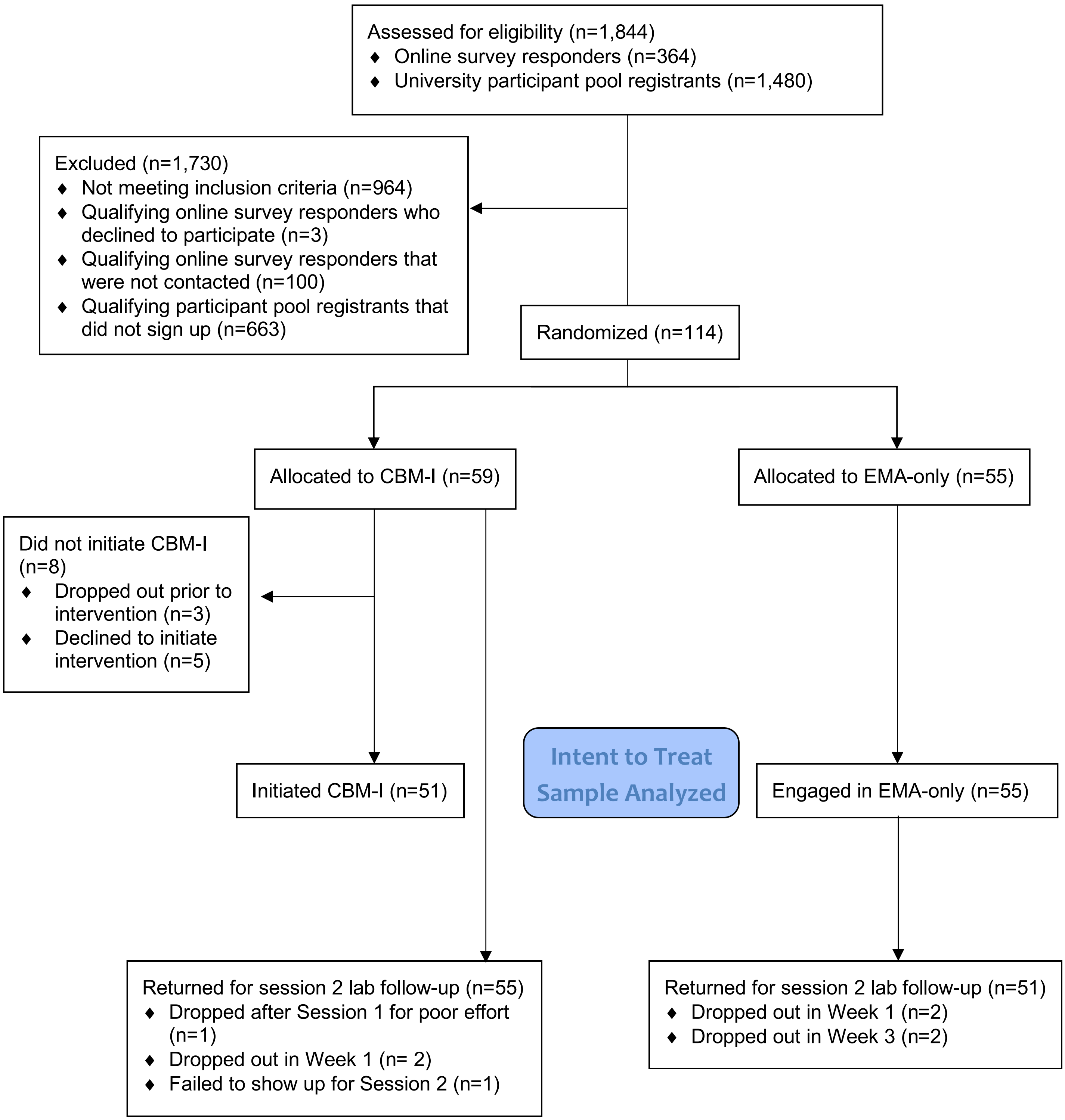
Yang, R., Cui, L., Li, F., Xiao, J., Zhang, Q., & Oei, T. P. S. (2017). Effects of cognitive bias modification training via smartphones. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2017.01370

Kuckertz, J.M. & Amir, N. (2014). "Cognitive Biases in Social Anxiety Disorder." In Social Anxiety: Clinical Developmental, and Social Perspectives, 3rd ed. Edited by S.G. Hofmann & P.M. DiBartolo, 483-510. New York: Academic Press.

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B. (2017). lmerTest package: Tests in linear mixed effects models. Journal of Statistical Software, 82(13), 1-26, doi: 10.108637/jss.v082.i13

Figures

Figure 1. CONSORT Diagram



Tables

Table 1. Outcomes of Interest

Cognitive Style Constructs	**Associated Trait Measures**	**Associated State Measures**
H1 Interpretation Bias	*H1a* Brief Body Sensations Interpretations Questionnaire (BBSIQ) – Negative Events Subscale	---
H2 Cognitive Reappraisal	*H2a* Cognitive Flexibility Inventory (CFI) - Alternatives Subscale *H2b* Emotion Regulation Questionnaire (ERQ) - Reappraisal Subscale	*H2c* EOD *self-rated ability to effectively use* cognitive reappraisal* *H2d* RT *likelihood* to use some form of emotion regulation during relatively high negative affect*
Symptom Constructs	**Associated Trait Measures**	**Associated State Measures**
H3 Social anxiety symptoms	*H3a* Social Interaction Anxiety Subscale (SIAS)**	*H3b* EOD reports of social avoidance* *H3c* RT reports of being calm vs. anxious*
H4 Fear of negative evaluation	*H4a* Brief Fear of Negative Evaluation Scale (BFNE) H4b. EOW past-week modified version of BFNE*	*H4c* EOD reports of being worried vs. comfortable with other people's thoughts of them*

Note: RT = Randomly Timed Survey; EOD = End of Day Survey; EOW = End of Week Survey.
*Assessed through ecological momentary assessment. All other measures were administered during the baseline lab session and 5 weeks later at the follow-up lab session.
**SIAS was used as a pre-screener to determine eligibility for enrollment, so there is variability in the amount of time between initial SIAS score and follow-up score.

Table 2. Sample Demographics

	CBM-I (*n* = 51)	EMA-only (*n* = 55)	Overall (*N* = 106)
Age	*M* = 20.67 *SD* = 2.92	*M* = 2024 *SD* = 3.12	*M* = 20.44 *SD* = 3.02
Gender			
Female	38 (74.5%)	40 (72.7%)	78 (73.6%)
Male	13 (25.5%)	15 (27.3%)	28 (26.4%)
Race			
White	36 (70.6%)	36 (64.7%)	71 (67.0%)
Asian	10 (19.6%)	7 (12.7%)	17 (16.0%)
Black	2 (3.9%)	5 (9.1%)	7 (6.6%)
Middle Eastern	0 (0%)	2 (3.6%)	2 (1.9%)
Multiracial	3 (5.9%)	5 (10.1%)	9 (8.5%)
Ethnic Identity			
Latinx/Hispanic	1 (2.0%)	2 (3.6%)	3 (2.8%)
Not Latinx/Hispanic	50 (98.0%)	52 (94.5%)	102 (96.2%)
Prefer not to answer	0 (0%)	1 (1.8%)	1 (0.9%)

Note. CBM-I = Cognitive bias modification for interpretations. EMA = Ecological momentary assessment.

Table 3

Model estimates for trait outcome models

Cognitive Style Outcomes	*Fixed Effect (b)*	*t*	*p*	Random Effect Variance *(SD)*	R^2M *(R^2C)*
Negative Interpretation Bias					.09 (*.66*)
(Intercept)	.23	1.62	.11	.58 (*.76*)	
Main Effect for Condition	.002	.01	.99		
Main Effect for Time	-.18	-1.44	.15		
Time by Condition Interaction	**-.57**	**-3.27**	**<.01**		
Cognitive Flexibility (CFI)					.01 (*.69*)
(Intercept)	.16	1.15	.25	.69 (*.83*)	
Main Effect for Condition	-.27	-1.34	.17		
Main Effect for Time	-.12	-1.07	.29		
Time by Condition Interaction	.09	.59	.56		
Cognitive Reappraisal (ERQ)					.03 (*.65*)
(Intercept)	.11	.79	.43	.63 (*.79*)	
Main Effect for Condition	-.35	-1.83	.07		
Main Effect for Time	.11	.89	.37		
Time by Condition Interaction	.09	.51	.61		
Social Anxiety Symptom Outcomes	*Fixed Effect (b)*	*t*	*p*	Random Effect Variance *(SD)*	R^2M *(R^2C)*
Social Interaction Anxiety Subscale					.15 (*.55*)
(Intercept)	**.37**	**2.93**	**<.01**	.41 (*.64*)	
Main Effect for Condition	.01	.08	.93		
Main Effect for Time	**-.79**	**-5.89**	**<.001**		
Time by Condition Interaction	-.02	.12	.90		
Fear of Negative Evaluation					.05 (*.71*)
(Intercept)	.20	1.50	.14	.67 (*.82*)	

Main Effect for Condition	-.004	.02	.98		
Main Effect for Time	**-.42**	**-3.92**	**<.001**		
Time by Condition Interaction	-.02	-.14	.89		
Weekly Fear of Negative Evaluation					.02 (*.79*)
(Intercept)	.13	.93	.35	.88 (*.94*)	
Main Effect for Condition	.04	.20	.84		
Main Effect for Time	-.21	-1.80	.08	.35 (*.59*)	
Time by Condition Interaction	-.16	-.96	.34		

Note. Beta estimates are partially standardized to allow for comparisons across studies. Significant effects are bolded. R^2M = Marginal R^2. R^2C = Conditional R^2. Marginal R^2 refers to the amount of variance explained by only the fixed effects in the model. Conditional R^2 refers to the amount of variance explained by the fixed and random effects in the model.

Table 4

Model estimates for state outcome models

Cognitive Style Outcomes	*Fixed Effect (b)*	*t*	*p*	Random Effect Variance *(SD)*	R^2M *(R^2C)*
Likelihood to use ER in high negative affect[+]					.02 (*.32*)
(Intercept)	-.65	-1.54	.12	1.14 (*1.07*)	
Main Effect for Condition	**.50**	**2.13**	**<.05**		
Main Effect for Time	.30	1.60	.11	.87 (*.93*)	
Time by Condition Interaction	-.06	-.24	.81		
Mean Affect per Person	.15	.54	.59		
Observations per Person	**.02**	**1.99**	**<.05**		
Daily likelihood to use cognitive reappraisal[+]					.01 (*.59*)
(Intercept)	**1.99**	**6.12**	**<.001**	3.83 (*1.96*)	
Main Effect for Condition	-.49	-1.09	.28		
Main Effect for Time	.38	.93	.35	2.91 (*1.71*)	
Time by Condition Interaction	-.43	-.85	.40		
Daily ability to use cognitive reappraisal					.01 (*.34*)
(Intercept)	.11	1.23	.21	.32 (*.56*)	
Main Effect for Condition	**-.27**	**-2.08**	**<.05**		
Main Effect for Time	-.11	-1.29	.20	.17 (*.41*)	
Time by Condition Interaction	**.27**	**2.10**	**<.05**		
Social Anxiety Symptom Outcomes	*Fixed Effect (b)*	*t*	*p*	Random Effect Variance *(SD)*	R^2M *(R^2C)*
State anxiety across all situations					.004 (*.35*)
(Intercept)	.11	1.47	.15	.30 (*.55*)	
Main Effect for Condition	-.02	-.17	.86		
Main Effect for Time	**-.16**	**-2.60**	**<.05**	.16 (*.39*)	
Time by Condition Interaction	.06	.69	.49		

State anxiety during social situations					.006 (*.35*)
(Intercept)	.15	1.91	.06	.32 (*.56*)	
Main Effect for Condition	-.03	-.26	.79		
Main Effect for Time	**-.19**	**-2.82**	**<.01**	.16 (*.40*)	
Time by Condition Interaction	.06	.60	.55		
Daily social concern					.005 (*.47*)
(Intercept)	-.13	-1.29	.20	.46 (*.68*)	
Main Effect for Condition	.07	.51	.61		
Main Effect for Time	**.13**	**2.15**	**<.05**	.06 (*.25*)	
Time by Condition Interaction	-.0004	-.004	.99		
Daily social avoidance					.01 (*.31*)
(Intercept)	.01	.15	.88	.28 (*.53*)	
Main Effect for Condition	.06	.56	.58		
Main Effect for Time	-.14	-1.81	.07	.13 (*.36*)	
Time by Condition Interaction	.12	1.17	.24		

Note. ER = emotion regulation. Significant effects are bolded. Unless otherwise noted, beta estimates are partially standardized to allow for comparisons across studies. + = beta estimates that are not partially standardized (due to binary outcome variable) and *z* values are reported instead of *t* values. R^2M = Marginal R^2. R^2C = Conditional R^2. Marginal R^2 refers to the amount of variance explained by only the fixed effects in the model. Conditional R^2 refers to the amount of variance explained by the fixed and random effects in the model.

Supplementary Material

Measuring emotion regulation strategy use.

To measure likelihood to use some form of emotion regulation, participants reported which of 19 different emotion regulation strategies they had engaged in over the 30-minutes before each RT survey prompt. Participants could select more than one strategy at a time, write-in additional strategies, or report they had not tried to change their thoughts and feelings over the previous 30 minutes. Participants were considered to have implemented a given strategy (i.e., received a “0” vs. a “1”) if they checked the box next to the layperson description of the strategy following the question, “Over the 30 minutes before the survey prompt, I tried to change my thoughts and feelings through….”

The strategies were: “ruminating about something” (rumination); “coming up with ideas/plans for action” (problem solving); “accepting them” (acceptance); “criticizing myself” (self-criticism); “thinking of the situation differently” (cognitive reappraisal); “thinking about the things that went/are going well” (thinking good thoughts); “pushing away bad thoughts” (thought suppression); “tackling the issue head on” (tackling the issue head on); “drinking alcohol” (alcohol); “using marijuana, nicotine, or other drugs” (drugs); “eating food” (eating); “exercising” (exercising), “TV/internet/gaming” (TV/gaming); “sleeping” (sleeping); “seeking advice/comfort from others” (advice-seeking); “ignoring/avoiding certain people/situations” (situational avoidance); hiding my thoughts/feelings from others (expression suppression); “doing something fun with others” (doing something fun with others). For each strategy above, participants saw the text that is enclosed in quotation marks, but not the conceptual labels included in parentheses.

Results of Specific Emotion Regulation Strategy Use Exploratory Analyses

Supplementary Table 1. Fixed effects for state outcome models assessing for change in the likelihood to report using each specific emotion regulation strategy during instances of personally high negative affect

Cognitive Emotion Regulation Strategies	*b*	*z*	*p*
Cognitive Reappraisal			
(Intercept)	**-3.13**	**-5.54**	**<.001**
Main Effect for Condition	-.19	-.70	.48
Main Effect for Time	-.48	-1.32	.19
Time by Condition Interaction	.14	.32	.75
Problem Solving			
(Intercept)	**-2.03**	**-4.16**	**<.001**
Main Effect for Condition	-.32	-1.31	.19
Main Effect for Time	.05	.28	.86
Time by Condition Interaction	**-1.07**	**-2.38**	**<.05**
Acceptance			
(Intercept)	-2.10	-1.75	.08
Main Effect for Condition	.07	.23	.82
Main Effect for Time	-.31	-.72	.47
Time by Condition Interaction	-.80	-1.36	.17
Rumination			
(Intercept)	-1.03	-1.72	.09
Main Effect for Condition	-.02	-.05	.96
Main Effect for Time	.07	.20	.84
Time by Condition Interaction	-.53	-1.23	.22
Thinking Good Thoughts			
(Intercept)	**-3.99**	**-5.78**	**<.001**

Main Effect for Condition	.02	.06	.95
Main Effect for Time	-.21	-.45	.66
Time by Condition Interaction	-1.12	-1.84	.07
<u>Self-Criticism</u>			
(Intercept)	**-2.56**	**-3.66**	**<.001**
Main Effect for Condition	-.34	-.88	.38
Main Effect for Time	-.004	.01	.99
Time by Condition Interaction	-1.08	-1.86	.06
<u>Thought Suppression</u>			
(Intercept)	**-1.72**	**-3.30**	**<.001**
Main Effect for Condition	.14	.43	.67
Main Effect for Time	-.20	-.67	.49
Time by Condition Interaction	-.09	-.28	.78
Behavioral Emotion Regulation Strategies	*b*	*z*	*p*
<u>Tackling the Issue Head On</u>			
(Intercept)	**-1.77**	**-4.42**	**<.001**
Main Effect for Condition	-.04	-.16	.88
Main Effect for Time	.22	1.17	.24
Time by Condition Interaction	.28	1.03	.30
<u>Distraction</u>			
(Intercept)	**-2.32**	**-4.73**	**<.001**
Main Effect for Condition	**.48**	**1.99**	**<.05**
Main Effect for Time	-.26	-.88	.38
Time by Condition Interaction	-.32	-.84	.40
<u>Watching TV/Using Internet</u>			
(Intercept)	-1.28	-1.23	.21

Main Effect for Condition	.15	.48	.63
Main Effect for Time	.11	.32	.75
Time by Condition Interaction	-.53	-1.11	.27
<u>Eating</u>			
(Intercept)	**-2.22**	**-2.98**	**<.01**
Main Effect for Condition	**.93**	**2.50**	**<.05**
Main Effect for Time	.24	.47	.64
Time by Condition Interaction	**-1.63**	**-2.85**	**<.01**
<u>Exercising</u>			
(Intercept)	**-5.73**	**-5.22**	**<.001**
Main Effect for Condition	.23	.38	.70
Main Effect for Time	.92	1.01	.31
Time by Condition Interaction	-1.62	-1.70	.09

Interpersonal Emotion Regulation Strategies	*b*	*z*	*p*
<u>Seeking Advice from Others</u>			
(Intercept)	**-3.88**	**-5.65**	**<.001**
Main Effect for Condition	-.34	-.99	.32
Main Effect for Time	-.17	-.37	.71
Time by Condition Interaction	.001	.002	.99
<u>Expressive Suppression</u>			
(Intercept)	**-4.58**	**-6.16**	**<.001**
Main Effect for Condition	-.04	-.12	.90
Main Effect for Time	-.44	-.77	.44
Time by Condition Interaction	-.47	-.71	.48

Note. All models controlled for the number of contributing observations and mean standardized negative affect for the first two weeks and the last two weeks of each participant's EMA responses. Time was entered as a random slope. Intercepts were allowed to vary. Multiple comparison corrections were not conducted and zero-inflated poisson distributions were used to account for the low frequency with which each strategy was reported. Results should be interpreted accordingly. Results from the following emotion regulation strategies are not included because models failed to converge due to low response rates: sleeping/taking a nap; using drugs; drinking alcohol; doing something fun with others; situational avoidance.

Significant interaction results interpreted.

Problem solving. The CBM-I group became less likely to report problem solving after the intervention than the EMA-only group

Distraction. The CBM-I group became more likely to report using distraction after the intervention than the EMA-only group.

Eating. The CBM-I group became less likely to report eating after the intervention than the EMA-only group.

Appendix A

The following deviations to the protocol occurred:

Administrative Details

Based on the expectation that there would be greater participant drop-out in the CBM-I (vs. EMA-only) condition, random assignment was originally weighted 60% to 40% in favor of the CBM-I condition. However, the anticipated asymmetry in actual participant dropout did not occur, so random assignment was re-weighted to 20% to 80% in favor of the EMA-only condition for the final 18 participants. This decision was made to balance participant numbers across conditions to allow for more robust statistical modeling.

Additionally, while MetricWire was typically programmed to deliver an End of Day (EOD) survey at 10pm every night, participants could request the survey be delivered earlier in the evening if they would be consistently unavailable at 10pm (e.g., they go to bed before 10pm). In these cases (n = 8), the EOD survey was reprogrammed to go off at either 8pm or 9pm for those particular participants.

Measurement Details.

Due to an administrative error, one item ("When mixing in a group, I find myself worrying I will be ignored") was mistakenly dropped from the post-intervention SIAS questionnaire. Thus, to assess for an intervention effect on social anxiety symptoms, the average item score across all available items in the pre-screen SIAS (20 items) and in the post-intervention follow-up SIAS (19 items) were compared. Additionally, given that the SIAS was used as a pre-screener to determine eligibility for enrollment, more than five weeks passed in between assessment points on this outcome measure. Further, due to scheduling constraints, there is variability in the amount of time between pre-screen and follow-up scores across participants.

Also, while the five CFI items that were included in the study represent those items that demonstrated at or above a .70 factor loading on the Alternatives Subscale, one additional item scoring above .70 ("When in difficult situations, I consider multiple options before deciding how to behave”) was not included in the data collection due to an administrative error.

Participant Payment.

Over the course of the study, the specific incentive schedule for the EMA monitoring portion was changed to increase participant compliance, although payment was always based on the proportion of completed surveys. Initially, n = 31 participants earned ~$0.20 for each completed EMA survey, and earned a bonus payment of $10 if they completed more than 80% of the EMA surveys, for a maximum payment of $60. Based on participant feedback, this flat-rate payment approach was changed to five payment brackets based on percentage of surveys completed, with a maximum payment possibility of $70. Specifically, participants earned $10 for completing 0-19% of surveys, $25 for completing 20-39% of the surveys, $40 for completing 40-59% of the surveys, $55 for completing 60-79% of the surveys, and $70 for completing 80-100% of the surveys. Based on additional participant feedback, and to increase participant motivation in the final weeks of EMA monitoring, the final n = 38 participants also received a $5 bonus for completing over 70% of EMA surveys during Week 4 and another $5 bonus for completing over 70% of EMA surveys during Week 5. Regardless of which version of the incentive schedule a participant was subject to, each of the six CBM-I sessions were compensated as equivalent to completing four EMA surveys.