Positive Imagery Training Increases Positive Self-Referent Cognition in Depression

Justin Dainer-Best[a,*], **Jason D. Shumake**[a,b], and **Christopher G. Beevers**[a,b]

[a]Department of Psychology, The University of Texas at Austin; 305 E. 23rd St.; Stop E9000; Austin, TX 78712

[b]Institute for Mental Health Research, The University of Texas at Austin

Abstract

Depressed adults often show a bias towards negative self-referent processing at the expense of positive self-referent processing. The current study assessed whether a mental imagery intervention (Positive Self Reference Training—PSRT) delivered via the Internet could improve self-referent processing and depressive symptomatology among adults with moderate or greater depression symptoms. Participants were recruited via online methods and randomly assigned to one of two computerized interventions: active PSRT ($n = 44$) or control training (NTC; $n = 43$). The PSRT involved visualizing the self in response to different positive cues (e.g., an achievement) every other day for two weeks. The NTC provided neutral cues about objects. Self-referential processing of positive and negative adjectives and depression symptoms were measured at baseline, one week, and two weeks after initiating training. Over those two weeks, PSRT participants showed a greater increase in positive self-referent processing than did NTC participants. Negative self-referent processing and symptoms of depression declined comparably in both groups. Similarly, for both groups, increase in positive and decrease in negative self-referent processing was associated with a greater reduction in depression. These results indicate that mental imagery has the potential to improve self-referential processing, especially for positive stimuli, which may, in turn, help reduce depressive symptomatology.

Keywords

depression; imagery; self-reference; SRET; online interventions

[*]Correspondence concerning this article should be addressed to Justin Dainer-Best, Department of Psychology, Bard College; Campus Road, PO Box 5000; Annandale-on-Hudson, NY 12504.

1. Introduction

Over the past half-century of depression-related research, the field has accumulated a great deal of evidence pointing to the role of cognitive biases in maintaining depressive symptoms (Beck, 1967; Gotlib & Joormann, 2010; Gotlib & Krasnoperova, 1998). The relationship of such biases and depressive disorders is well established; individuals diagnosed with depressive disorders are more likely to have negative self-schemas and to view themselves negatively (Beck, 1967). Beck theorized that negative views about the self, the world, and the future (the cognitive triad) contribute to the maintenance of depression.

Such negativity is mirrored by a difficulty relating to positive information; Beck referred to this as a "positive blockade" (1967). Whereas negative self-referent information is preferentially processed with relative ease, positive information tends to be ignored, discounted, or processed with difficulty. Together, these biases help to define the existing memory representations of the world, including self-representations. Individuals with negative self-views (or schemas) are more likely to experience increased depressive symptomatology (Connolly et al., 2015; LeMoult et al., 2017b).

The relationship between negative self-schema and the maintenance of depression is sometimes viewed in causal terms, with stronger negative self-schemas thought to produce a more protracted episode of depression (Beck, 1967). It is clear that both positive and negative schema are strongly correlated with depressive symptoms (Phillips et al., 2010; Dainer-Best et al., 2018), and can predict worsening of symptoms (Disner et al., 2016); however, such studies are correlational and thus vulnerable to third variable explanations that account for the association between self-referent processing and depression.

As such, attempts to modify self-schema can be viewed as a direct test of whether maladaptive cognitions serve to maintain depression, as reductions in self-schema should precede reductions in depression symptoms, if there is a causal link. Although cognitive-behavioral therapy (CBT; Beck, 1979) targets maladaptive cognitions, few studies to date have directly measured self-referent information processing, with assessments such as the Self-Referent Encoding Task (SRET; Derry & Kuiper, 1981), before and after CBT. While schema change is a hypothesized mechanism for depression improvement, randomized controlled trials rarely track biased self-referent processing over the course of treatment.

Two such studies of adults with social anxiety disorder (Goldin et al., 2013; Thurston et al., 2017) measured self-referent processing with the SRET before and after CBT. Both studies found reduced negative self-reference and increased positive self-reference following treatment. With a depressed sample, Quilty et al. (2014) measured change in self-reference in pharmacotherapy and CBT with the SRET. They reported that self-referent processing changed similarly across both treatment conditions. Further, Dozois et al. (2009) reported that depressed adults who received cognitive therapy in conjunction with pharmacotherapy saw improvement in positive and negative self-schemas that was not observed in people who received pharmacotherapy alone. Both treatment groups reported similar reductions in depression during treatment.

CBT interventions are multifaceted, and so even if change in self-referent processing is observed, it is unclear what aspect of treatment may be producing those changes. In contrast, there is a nascent literature using cognitive bias modification (CBM) techniques to target and change specific cognitive biases associated with depressive disorders (Hallion & Ruscio, 2011; MacLeod et al., 2009). Studies that target specific cognitive mechanisms with CBM have generally been focused on attention (Wells & Beevers, 2010; Ferrari et al., 2016), overgeneral memory (Watkins et al., 2009; Raes et al., 2009; Neshat-Doost et al., 2013), and interpretation (Holmes et al., 2009; Joormann et al., 2015; LeMoult et al., 2017a). Assessing whether such focused interventions leads to symptom reduction is ideal for theory testing, and continues to be an important focus of clinical research (MacLeod et al., 2009).

Interpretation-focused CBM studies in particular have used imagery-based techniques to remediate maladaptive cognitive processes (Holmes et al., 2009; Hitchcock et al., 2016). Such techniques rely on participants' mental imagery rather than providing cues to specific stimuli. They may also have a more idiographic target, focusing on individualized information processing rather than directing all participants to specifically attend to the same stimuli. Such cue-based mental imagery interventions may directly target maladaptive schema in depression (Holmes et al., 2016) and elicit emotional reactions as though imagined events were happening (Holmes & Mathews, 2010).

Mental imagery interventions that focus on depression-specific maladaptive cognitions have the potential to change self-referent cognitions. Given a biased self-schema, any such interventions must change the way that judgments of information about the self are made. Previous research has shown that making self-referential judgments about emotional material enhances specific verbal and visual information and can boost memory for that information (Hamami et al., 2011 Thus, by repeatedly encoding positive self-referent information through mental imagery, it becomes more salient when self-schema is accessed. This prioritization makes a positive self-schema more accessible, thus reducing negative self-referent processing biases. Studies have shown that positive memory activation can improve mood following simple procedures (Hall et al., 2018), which may have an overall impact on mood and self-schema.

However, memories are (by definition) limited to accessible retrievals—given the theoretical (Beck, 1967) and empirical (Bradley et al., 1995; Gupta & Kar, 2012 memory biases seen in depression, imagery techniques may focus instead on a future orientation. Klein (2013) has additionally theorized that future-oriented imagery is straightforward even when individuals have no episodic memory of a situation. Other work has also shown that simulating positive future events in a single session can increase depressed participants' ratings of the likelihood and importance of other future events (Boland et al., 2017).

Given the robust association between the self-referential bias and dysphoria, we developed a CBM intervention focused on enhancing positive self-referent processing by encouraging participants to repeatedly practice viewing themselves positively and imagining a pleasant and positive future. To provide an adequate control, we developed a neutral intervention that focused instead on detailing objects in a neutral, present situation. This training directly mirrored the positive intervention, but without what we believed to be the active components

of that intervention. That is, participants did not focus on positive, self-referent, or future-focused imagery.

We thus compared a positive self-reference training (PSRT), focused on using positive cues to encourage positive, future, self-referent processing, to a neutral training condition (NTC) that used neutral cues to encourage neutral, current, non-self-referent processing. We hypothesized that participants in the PSRT—but not the NTC—group would show improvement on measures of self-referential bias (i.e., the SRET), especially in terms of their response to positive items. We also hypothesized that the reduction in depressive symptoms in the PSRT group would be linked to their improvement in self-reference.

2. Methods

2.1. Participants

Participants in the study were eligible if they were adults between the ages of 18-45, had elevated total score > 13 on the Center for Epidemiologic Studies–Depression (CES-D) scale, were fluent in English, and provided informed consent. Participants who completed the study (completers; $N = 87$) were on average 26.4 years old ($SD = 7.0$), female (84%), and white (66%). They were paid between $25 and $35 for completing the study, with higher payments being received for completing additional training sessions as discussed below. Participants in the intention-to-treat (ITT) sample, who completed the full baseline assessment and thus were randomized into one of the two training conditions ($N = 264$), were on average 26.8 years old ($SD = 7.0$), female (75%), and white (67%).

Participants were recruited through online postings advertising an online mood study using three online forums: (1) Craigslist, an online bulletin board, in several major cities; (2) ResearchMatch, a national health volunteer registry supported by the U.S. National Institutes of Health as part of the Clinical Translational Science Award (CTSA) program; and (3) online postings through a community events board at the University of Texas at Austin. The advertisements directed participants to a website to determine study eligibility. Here, participants provided informed consent, provided their age, and filled out the CES-D. If they were eligible for the study, further measures (discussed below) were collected, and participants were randomized into one of two training conditions.

2.2. Measures

2.2.1. Center for Epidemiologic Studies–Depression Scale (CES-D)—The CES-D (Radloff, 1977) is used to assess depressive symptoms over the past week using a 20-item self-report questionnaire. Scores may range from 0 to 60, with higher scores indicating elevated depressive symptoms. Scores greater than 16 have been regularly used as an indication of possible diagnosis of major depressive disorder (Radloff, 1977; Santor et al., 1995); scores above 13 may thus indicate mild or greater depressive symptoms. The CES-D was used to determine eligibility at baseline, and assessed again following one week (T1) and two weeks (T2).

2.2.2. Self-Referent Encoding Task (SRET)—The SRET (Derry & Kuiper, 1981) is an affective decision-making task designed to assess schema-related processing. Participants

make decisions about whether positive and negative adjectives are self-descriptive. Words appear on their computer screens, and participants make rapid judgments following each word's display. The SRET was presented using Inquisit software (Millisecond Software LLC, Seattle, WA). Inquisit is a commercial platform that runs "full screen" on participant's computers as a Java applet. The SRET was administered at baseline (T0), one week (T1), and two weeks (T2) after baseline.

The SRET consisted of a short practice block followed by the actual task. The practice block allowed participants to view and respond to 5 practice words. After completing the practice block, 90 words were displayed once each in random order. Words were displayed in white text on a black screen and remained on-screen until participants responded. Participants were told to use the Q or P keys on their personal keyboard to answer whether the word described them or not. Each trial was followed by a 1,500 ms intertrial interval. Words were chosen from among the words used in previous SRET studies (Disner et al., 2016; Auerbach et al., 2015; Dainer-Best et al., 2017) and had been selected from the Affective Norms for English Words (Bradley & Lang, 2010) and from an additional validated word-list (Doost et al., 1999).

As the SRET was presented three times (at T0, T1, and T2), a subset of words were presented only once. Thus, 80 words were presented every time,[1] and an additional 30 words, chosen randomly from all possible words, were presented once each: 10 at baseline, 10 at T1, and 10 at T2.[2] Although free recall data has often been collected in past work, we did not collect such data in this study due to concerns about learning effects due to repeated administration. Cronbach's alpha was calculated for each valence for each time-point, with 2,500 bootstrapped iterations to calculate 95% confidence intervals. Lower bounds exceeded .92 for alpha across all three time-points for positive responses on the SRET, and .90 for negative words. We also calculated Cronbach's alpha on change scores (i.e., T0 – T2 scores for each core word), with 2,500 bootstrapped confidence intervals. For positive words, alpha = .86, 95% CI [.81, .89], while for negative words, alpha = .90, 95% CI [.87, .93]. Automated procedures did not single out any word for substantially raising or lowering alpha by exclusion from the scale.

The primary behavioral outcome from the SRET is the probability of endorsing positive vs. negative words, which we have demonstrated to be strongly linked to depression (Dainer-Best et al., 2018). Additionally, including an extra subset of words at each time-point allowed us to distinguish between responses that varied due to repetition and those that varied putatively due to changes in self-referential processing.

---

[1] The following positive words were shown each time: joyful, gentle, cool, funny, surprised, glad, loyal, bold, proud, good, outstanding, awesome, elated, adorable, admired, devoted, untroubled, festive, content, beautiful, satisfied, thoughtful, free, excellent, dignified, confident, excited, kind, capable, brilliant, best, helpful, nice, alive, pleased, terrific, bright, vigorous, carefree, playful. The following negative words were shown each time: disloyal, alone, dumb, hostile, helpless, worried, defeated, annoyed, obnoxious, mad, shamed, morbid, anguished, unhappy, guilty, distressed, frustrated, ashamed, weak, disgusted, wicked, brutal, unloved, displeased, depressed, sad, terrified, fearful, violent, burdened, lonely, crushed, angry, scared, bad, hateful, upset, troubled, afraid, sorry.
[2] These words included, at baseline: wonderful, lucky, useful, fantastic, grateful (positive) and useless, cruel, nasty, terrible, dreadful (negative); at T1: great, fun, jolly, loved, engaged (positive) and rejected, lost, stupid, insane, unwanted (negative); and at T2: cute, hopeful, friendly, famous, masterful (positive) and insecure, sinful, horrible, rude, bored (negative).

2.2.3. Training—There were two training conditions, positive self-reference training (PSRT) and a neutral training condition (NTC). The PSRT prompted participants to focus on (1) future events that were (2) specific, (3) positive or fun, and (4) involved the participants' involvement. Positive cues focused on places or concepts that could engage self-referent positive imagery, like going to a café or receiving a gift; instructions explicitly suggested scenarios that would be "positive, pleasant, or fun".[3] The NTC, conversely, was designed to lack the primary components that were active in the PSRT condition. As such, the NTC training instructed participants to focus on (1) present locations that were (2) generic, (3) neutral, and (4) involved primarily objects and spaces. Neutral cues were all places. The practice cues were a café and a gift shop.[4]

Both training conditions were couched in the idea of "target[ting] unhelpful ways of thinking". Participants viewed short (two minute) videos elaborating the details of their training; the full text of each training video is included in the supplementary materials.[5] They were reminded of the important details of their training condition before completing the training on each day that they completed it. The first day, they were also emailed a link to view the training video again. Both trainings provided participants, one at a time, with two cues per day, and asked them to create audio recordings of themselves responding to the cue by imagining the scenario and describing it according to the above, numbered foci. Participants were not given instructions as to how long they should spend imagining the scenario before recording, but were told to try to verbalize their response for 3–5 minutes. The prompt and cues both differed by condition, although there was some overlap in cues. Two cues were given each day of the training; the first day, these were described as "practice cues".

Following the video training—and every subsequent day that they responded— participants were instructed to complete the cue recordings for their condition, talking into a microphone about the cue for 3–5 minutes. Thus, a cued imagery training session could be done in as quickly as six minutes, although some participants completed longer recordings and some shorter. Participants received emailed reminders to complete a minimum of four and a maximum of seven additional pairs of recordings, every other day. Thus, participants who completed the study submitted between 10 and 16 cued recordings, practice recordings included. (Participants who submitted fewer than ten recordings were not considered adherent, as they had not completed the study.) Completing more recordings was encouraged with a hierarchical payment structure. Participants were paid $25 if they completed all assessments and a minimum of ten recordings. For each additional set of recordings that participants completed (i.e., 12, 14, or 16), participants were paid an additional amount ($2, $3, and $5), enabling participants to earn a maximum of $35. This information was included in the consent form and reiterated after randomization.

---

[3]The remaining PSRT cues were: going shopping, an achievement, feeling relaxed and comfortable, a celebration, being in a peaceful place, playing a game, feeling capable, listening to music, being somewhere beautiful, being extremely interested in something, being supportive, a garden or museum, a great restaurant, your favorite movie.
[4]The remaining NTC cues were: a clothing store; a theatre; a bedroom; a banquet hall; an office; a game room; a classroom; a music hall; a palace; a science museum; a restaurant; an art museum; a garden; a meeting room.
[5]Supplementary materials, including scripts for both conditions, analysis files, and data, are available on the Texas Data Repository, https://doi.org/10.18738/T8/RHEMGW.

All recordings were voice-only recordings spoken directly into a microphone, on participants' personal phones or computers. Participants were able to complete the recordings either through a "chat bot" in Telegram, a mobile application available for major smartphone platforms (Telegram Messenger LLP, Berlin, Germany), or through a website using a jQuery (jQuery Team) plugin to record from a computer microphone. Both methods uploaded participants' voice recordings to a secure website to ensure training adherence. Of those who completed the study, in the NTC group, 31 participants used Telegram and 12 participants used the web interface; in the PSRT condition, 25 participants used Telegram and 19 participants used the web interface. Following every day's cue recordings, participants were asked to rate how well they believed they had been able to imagine the scenario they were recounting.

2.2.4. Assessment of subjective benefit and acceptance—At the end of the study, before completing other measures, participants completed a ten-item questionnaire assessing their subjective satisfaction with the training program. These questions were drawn from those reported in studies of online interventions (Meyer et al., 2009), modified to best fit the current work.

2.3. Procedure

The Institutional Review Board at the University of Texas at Austin approved all procedures. The trial was preregistered prior to beginning data collection at ClinicalTrials.gov (Identifier NCT03056963; Dainer-Best, 2017). We proposed to recruit 40 participants in each condition, following previously published studies (LeMoult et al., 2017a; Hitchcock et al., 2016; Boland et al., 2017).

Participants provided signed informed consent online using REDCap electronic data capture tools hosted at UT-Austin (Harris et al., 2009). REDCap (Research Electronic Data Capture) also presented all questionnaires. Following the completion of CES-D screening, demographic information were collected. Participants were asked to self-define a personal identification phrase that was specific but did not contain protected health information (PHI). This phrase was used to link data between training recordings, self-report questionnaires, and the behavioral task. They were then automatically directed to a separate website to complete the SRET. The SRET took between eight and 10 minutes.

After completing the SRET, participants were randomized into either a positive (PSRT) or neutral (NTC) condition. Randomization was not stratified in any way and was entirely random for the majority of the data collection period (i.e., three months). Randomization occurred without human action through URL redirection; participants landed upon a website which was randomized to point to one of two conditions. Near the end of data collection (i.e., for one week), once desired sample sizes had been met, in order to conclude the study with an approximately equal number of completers between intervention groups, participants were randomized at a rate of 4:1 to the neutral condition, again through URL redirection. Seven NTC participants and one PSRT participant completed the study during this 4:1 allocation period. Participants were able to enroll beginning on March 17, 2017; the final participant concluded their participation on June 21, 2017.

Once randomized, participants were again provided with information about payment for the study as described above, and asked to confirm that they understood this information. Participants were blind to their condition, and interaction with experimenters was minimal. They were then provided with a short video explaining the training for their condition, as described above. They completed a practice set of cues. Every other day henceforth and for the following two weeks, at the same time as they finished watching the training video, each participant received an email reminding them to log in to REDCap to complete the day's training. Thus, participants received a reminder email for seven trainings. On the fourth day of training (i.e., eight days following baseline) participants completed the CES-D and SRET, before being invited to complete that day's training. This schedule can be seen in Figure 1.

After all trainings had been completed, on the 15th day following baseline, participants received an email asking them to once more complete the CES-D and SRET; they also completed the subjective benefit questionnaire at this time. Participants were paid $25 to $35, as described above. All study procedures were automated, although research personnel were available to troubleshoot any technical issues that occurred through data collection.

2.4. Participant Attrition and Data Filtering

The screening was begun by 1,243 participants, of which 678 (54%) were eligible and completed the remaining baseline questionnaires. (See Figure 2 for a CONSORT diagram (Moher et al.) describing participant flow.) Only 264 participants were randomized—the remaining participants were not eligible, did not download the software to complete the SRET, or did not provide their identification phrase. Of participants who were randomized to the PSRT condition ($N = 124$), 31% completed the study (i.e., four or more training sessions as well as baseline, one-week [T1], and two-week [T2] assessments). Of participants who were randomized to the NTC condition ($N = 140$), 35% completed the study after randomization. We refer to these participants as completers. Eight participants completed the baseline and T2 assessments, but not T1, and were not included as completers; only three of these participants also were adherent to completing the cued trainings. All participants who completed all three assessments were adherent to the training. Most participants completed either all 16 possible recordings ($N = 43$) or 14 recordings ($N = 19$); a minority completed only 12 recordings ($N = 9$) or 10 ($N = 16$).

Participants' questionnaire data was dropped for a specific time-point if all questions on the questionnaire were answered the same despite some items being reverse scored or if "catch" items (e.g., directing participants to select a certain answer to a question) were not answered correctly ($N = 10$ [3.8%] at baseline, 1 [1.0%] at T1, and 2 [2.3%] at T2). These metrics were taken as an indication that participants were likely not attending to the self-report questionnaires, and thus individual responses were dropped.

On the SRET, trials were marked as outliers if reaction times (RTs) were less than 200 ms (0.9% of all trials). These trials were dropped from analyses. After dropping trials based on RT, those participants with more than 25% of trials deemed outliers for any reason were excluded for that time-point ($N = 3$ [3.4% of completers]).

2.5.   Analysis of Missing Data

As described above, a large percentage of data were missing owing mainly to participant attrition and, to a lesser extent, data cleaning. To help diagnose the mechanism of missing response data, a logistic regression predicting dropout as a function of treatment condition and baseline CES-D found no interaction between the two, $\beta = 0.01$, 95% CI [−0.04, 0.06], $p = 0.59$. There was a near-significant effect of CES-D that was independent of treatment assignment, $OR = 1.02$, 95% CI [1.00, 1.05], $p = .07$, with greater baseline depression increasing the likelihood of dropping out. (Median CES-D for completers = 28, interquartile range 19–34; median CES-D for dropouts = 31, interquartile range 21–39.) There was no effect of condition, $OR = 0.93$, 95% CI [0.55, 1.59], $p = .79$. This strongly suggests that data were missing due to similar factors in both groups (i.e., somewhat higher dropout among more severely depressed participants), not the type of training offered.

Lastly, method of recording (Telegram vs. web) did not predict dropout. There was no interaction between method and condition, $\beta = −0.77$, 95% CI [−1.87, 0.32], $p = 0.17$, nor was there a significant main effect of method, $OR = 0.83$, 95% CI [0.49, 1.39], $p = .48$.

2.6.   Data Analytic Plan

Data cleaning, modeling, and visualization were conducted in RStudio (version 1.0.136) running R (version 3.3.2) with the following packages: *dplyr* (Wickham et al., 2015), *tidyr* (Wickham, 2017), *ggplot2* (Wickham, 2009), *effects* (Fox, 2003), and *lme4* (Bates et al., 2015).

Because we were most interested in the effect of the treatment protocol on negative self-referent processing, primary analyses focus on the completer sample. However, to assess the impact of non-compliance and missing outcomes on the intervention's efficacy, intention-to-treat analyses are also performed. Notably, non-compliance is particularly an issue with online intervention studies, as it requires relatively little effort on the part of the participant to complete baseline measures and be randomized into a treatment arm. Indeed, many participants may start an internet intervention but then drop out quickly due to time constraints, lack of motivation, technical problems, lack of interpersonal contact, perceptions of intervention ineffectiveness, among other reasons (e.g., Christensen et al., 2009).

Primary analyses were conducted using generalized linear mixed models that predicted trial-level binomial response (the log-odds of endorsing an SRET word as self-descriptive) using predictors of condition (PSRT, NCT), time-point (baseline, T1, T2), valence (positive, negative), and all possible interaction terms. Random effects were modeled to allow random intercepts and slopes over time-point by participant, word, and word novelty. These primary analyses focused on the completer sample (i.e., participants who completed the baseline, T1, and T2 assessments). We also repeated analyses with the intention-to-treat sample (ITT; i.e., any participants who were randomized to PSRT or NTC and provided their ID after randomization, as per Hollis & Campbell, 1999) using the same generalized mixed-effects logistic regression models as in the completer analysis. Analyses can be reviewed in supplementary materials accompanying this article.

Secondary analyses were designed to test whether the change in SRET for the PSRT group was more strongly linked to change in CES-D as compared to the NTC group. These exploratory analyses used a simultaneous latent difference score (LDS) model, also referred to as a bivariate LDS model, a structural equation modeling (SEM) approach to the analysis of multivariate change.[6] LDS models emphasize how individual differences in change in one variable predict individual differences in change in another variable (Newsom, 2015). Models were created using *OpenMx* in R (Neale et al., 2016; Pritikin et al., 2015; Boker et al., 2017). These analyses can also be reviewed in the supplementary materials.

Because this model is based on the covariance both within and between the two instruments (the SRET and CES-D), it is important that the same items be used for both instruments at all measurement occasions and that scores for both instruments are similarly scaled. Therefore, we did not include the novel SRET words in this analysis. Per participant, we only included words on which the participant had valid responses at all three timepoints. We summed numbers of endorsed words per valence, per-time-point. We corrected for trials lost by creating a per-person average multiplied by the total number of possible trials (e.g., (number of positive endorsements number of positive trials at T1) × 40). A maximum of four trials was lost for any one time-point. To create a composite difference score for the SRET, we subtracted the number of positive words endorsed from the number of negative words endorsed for each time-point, and then subtracted the minimum of all scores across all time-points so that the measurement scale would be comparable to that of the CES-D in terms of range and variance. In order to diminish the impact of extreme values and achieve multivariate normality, scores greater than 3 median absolute deviations from the median were proportionally reduced to eliminate skew while preserving the original rank order. This was applied to six CES-D scores at T0, two at T1, and three at T2. For example, the two highest scores, 51 and 56, were reduced to 47 and 48, while remaining the highest scores in the distribution.

In LDS models, a latent variable representing the difference score is specified for each measurement at T1 and T2, and the autoregressive path between each measurement is set equal to 1. We initially fit three major models:

---

[6]Some readers may wonder why we chose to use this approach over a more conventional mediation analysis, in which the treatment condition's effect on CES-D was mediated by its effect on SRET. There were several reasons for this decision. First, there was no evidence that the two training conditions had a differential effect on depression outcome, and testing for mediation effects in the absence of a significant total effect, while appropriate in some circumstances, can severely elevate the risk of Type I error (Loeys et al., 2015). Second, the model that showed a significant effect of training condition on self-schema was a generalized logistic mixed effects model of self-schema as a group × valence interaction on the probability of individual word endorsement, controlling for the random effects of the individual words and whether they were familiar or novel at each time point. Directly combining this model with the linear model of total CES-D score into a mediation analysis is not a simple matter because the error family, link functions, and random effects structures are incompatible. Therefore, a transformation of the individual word endorsements into a total score, similar to what we did for the LDS model, would be needed to make a mediation analysis work. But, third, when self-schema was operationalized in this reduced way (ignoring the trial-level effects of specific word stimuli) there was no longer a detectable effect of treatment on SRET outcome. Without indication of treatment affecting either the main outcome (CES-D) or the purported mediator (SRET), we saw no utility in building a causal mediation model relating two null effects. The simultaneous LDS model is different from a mediation model in that it accounts for individual trajectories, the temporal dynamics of which could theoretically differ between groups in the absence of mean differences. For example, one individual could show a large improvement in the SRET but no improvement in CES-D, while another individual could show no improvement in the SRET but a large improvement in CES-D. But when these two individuals are averaged together, both SRET and CES-D will show parallel trends of modest improvement even though, for these individuals, SRET change is in fact negatively correlated with CES-D change! Inspection of the individual trajectories plotted in Figures 3 and 4 reveals a great deal of complexity and diversity in how individuals change from week to week. We believe an attempt to characterize and quantify this aspect of individual change is important.

1. A baseline no-change and no-group-differences independence model, that estimated only two means and two variances (one for SRET and one for CES-D), which were independent of each other, constant over time, and equivalent between groups.

2. A no-group-differences model, which, in addition to the T0 mean and variance of SRET and CES-D, estimates all of the following parameters as being equal between groups: the covariance between SRET and CES-D at baseline, the mean and variance of latent change in SRET and CES-D per week, the covariance among T0 and latent change measurements within each measure, and the simultaneous (lag-0) effect of SRET change on CES-D.

3. A model which includes all of the above parameters but estimated separately for the PSRT and NTC groups. A final model was then derived by setting specific parameters to be equal between groups if their difference did not lead to a significant improvement in fit. Models were fit to the raw data using full information maximum likelihood (FIML) estimation.

We also evaluated the possibility of a lagged (non-simultaneous) effect of SRET change on CES-D change, in which T1 SRET change predicts T2 CES-D change, but the addition of this path was not significant, $\chi^2(1) = 0.03$, $p = .85$.

To address concerns about whether we had an adequate sample size for this type of analysis, we conducted a Monte Carlo simulation to estimate its power to detect the primary effect of interest, namely, a group difference in the path coefficient from SRET to CESD. The simulation indicated that our sample size should have given us an 87.9% chance of detecting this difference at $p < .05$, if the hypothesized effect were true. These analyses can also be reviewed in the supplementary materials accompanying this article.

3. Results

3.1. Baseline Summary Statistics

Participants in the PSRT and NTC completer groups did not differ on age ($t(80.7) = 1.32$, $p = .19$), sex ($\chi^2(1) = 0.06$, $p = .81$), gender ($\chi^2(2) = 1.61$, $p = .45$), race ($\chi^2(4) = 2.80$, $p = .59$), ethnicity ($\chi^2(1) = 0.006$, $p = .94$), self-reported current major depressive episode ($\chi^2(1) = 0.007$, $p = .93$), or depression severity as measured by the CES-D ($t(84.4) = .88$, $p = .38$). Both samples had elevated depression symptoms, median CES-D = 28.5, interquartile range 19–34. Demographic information is presented in Table 1, as are endorsements and reaction times (RTs) from baseline on the SRET. A logistic mixed-effects model with factors of group (PSRT or NTC [reference level]) and word valence (positive or negative [reference level]) predicting response at baseline found no significant interaction, $\beta = -0.06$, 95% CI [−0.28, 0.14], $SE = 0.11$, $p = .52$. Without the interaction, there was neither a significant main effect of valence ($OR = 1.00$, 95% CI [0.61, 1.62], $p = .98$) nor of condition ($OR = 0.90$, 95% CI [0.67, 1.23], $p = .52$), indicating that the treatment groups did not differ at baseline in terms of their SRET endorsements.

### 3.2.	Effect of Training on Self-Referent Processing

To test whether the PSRT or NTC group membership changed self-schema as measured by the SRET, we defined generalized linear mixed models as described in the Data Analytic Plan above, modeling endorsement of positive or negative words as a logistic model. Wald tests in completers indicated a significant interaction between group, valence, and a linear trend over time, $\beta = 0.33$, 95% CI [0.12, 0.55], $SE = 0.11$, $p = .003$. A likelihood ratio test also indicated statistical significance, $p = .01$, with the addition of the three-way interaction, when compared to a model without it. A parametric bootstrap resampling of the data with 2,500 simulations of the likelihood ratio test, conducted using the *pbkrtest* package (Halekoh & Højsgaard, 2014) in R, was also significant, $p = .01$.

To facilitate interpretation of the interaction, we refit the simple model of endorsement as a function of time (using the same random effects structure as the full model) within each group × valence cell, and computed the ratios of post-treatment endorsement odds over pre-treatment endorsement odds. The PSRT group experienced a significant reduction in the odds of endorsing a negative word, $OR = 0.46$, 95% CI [0.40, 0.53], $p < .001$, as did the NTC group, $OR = 0.51$, 95% CI [0.44, 0.58], $p < .001$. Given that the OR for each group lies within the 95% CI of the other, the magnitude of improvement was not meaningfully different between groups ($p = .22$). In contrast, the PSRT group showed a significant increase in the odds of endorsing a positive word, $OR = 1.59$, 95% CI [1.40, 1.80], $p < .001$, as did the NTC group, $OR = 1.18$, 95% CI [1.04, 1.35], $p = .01$. However, unlike negative endorsements, the ORs and 95% CIs for positive endorsements indicate a meaningfully greater benefit of PSRT over NTC ($p = .002$). Thus, the 3-way interaction can be interpreted as a proportionally greater increase in positive endorsements for the PSRT group with a more equivalent decrease in negative endorsements for both PSRT and NTC groups. This comparison can be seen in Figure 3. Visualizations of change of responses to individual words across time, between groups, can be viewed in the supplementary materials.

### 3.3.	Effect of Training on Depression Symptom Change

To test whether the PSRT or NTC conditions reduced self-reported depression symptoms, we ran a mixed-effects linear regression predicting CES-D score, with predictors of training condition and a linear ordered factor of time. The interaction of condition × time was not significant, $\beta = 0.85$, 95% CI [−1.91, 3.61], $p = .55$. Without the interaction, there was no main effect of condition, $\beta = −0.96$, 95% CI [−4.90, 2.97], $p = .63$. There was a significant main effect of time ($p < .001$), with an improvement of 7.3 points, 95% CI [5.4, 9.2], in CES-D score over the course of 2 weeks. As seen in Figure 4, across both groups, CES-D scores reduced in a linear trend from baseline ($M = 28.7$; $SD = 10.8$) to T2 ($M = 21.4$; $SD = 10.2$), a 25% reduction.

### 3.4.	Latent Difference Score (LDS) Models Relating SRET Change and CES-D Change

To clarify the relationship between schema and depression change, we built a two-group simultaneous LDS model for the PSRT and NTC groups, as described in the Data Analytic Plan above. We tested hypotheses about the equality of group dynamics by evaluating the likelihood ratios of models fit with and without equality constraints in the group parameters. For this analysis, we needed to construct a composite measure of self-schema negativity that

could be treated as a continuous time-varying covariate of CES-D scores. Given that the probability of negative endorsements decreased over time and the probability of positive endorsements increased for both groups, we operationalized aggregate SRET negativity as the difference between total negative endorsements and total positive endorsements, with a unit increase in this score corresponding to either one more negative, or one less positive, endorsement. We hypothesized that change in the SRET would predict change in CES-D more strongly in the PSRT group than in the NTC group.

The LDS model that assumed group equivalence for all parameters included 14 parameters and was a large improvement over the baseline no-change model, $\chi^2(10) = 236.1$, $p < .001$. However, the overall fit of this model was not good: it was significantly different from an exact fit, $\chi^2(40) = 68.4$, $p = .003$, and had a Root Mean Square Error of Approximation (RMSEA) equal to 0.10, 90% CI [0.06, 0.13]. Thus, we then conducted an omnibus test of group differences by fitting the same model while allowing the groups to differ on every model parameter, thus doubling the number of estimated parameters to 28.

This model led to a further significant improvement in fit, $\chi^2(14) = 32.0$, $p = .004$. The overall fit of this model was better, $\chi^2(26) = 36.4$, $p = .09$, RMSEA = 0.07, 90% CI [0.0, 0.12], but still not good. We subsequently fit a series of models that imposed group-equality constraints on the model parameters one at a time. Those constraints that led to a significant loss in model fit were rejected. This resulted in a final model in which the groups were allowed to differ on only four parameters: variance in SRET change and disturbance in CES-D change, covariance between T1 and T2 SRET change, and covariance between T1 and T2 CES-D change. Overall, the fit of this model was good, $\chi^2(36) = 45.2$, $p = 0.14$, RMSEA = 0.06, 95% CI [0.0, 0.11]. However, the wide confidence interval surrounding the RMSEA estimate suggests the model might benefit from a larger sample size.

The final latent difference score models can be seen in Figure 5. Coefficients reported here and in the figure are in standardized correlational units (i.e., −1 and 1 represent perfect negative and positive linear relationships), and Table 2 reports the unstandardized coefficients and their standard errors, along with the standardized coefficients (if applicable), and the χ^2 and p values from the likelihood ratio test comparing models in which that parameter was either fixed to equality between groups or free to vary between groups. Notably, the path influence of SRET change on CES-D change was not significantly different between groups, $\chi^2(1) = 0.1$, $p = .73$, suggesting that the PSRT intervention did not alter the fundamental relationship between self-schema and depression. However, the dynamics governing within-person change appeared to differ markedly between training conditions. This was evidenced by significantly less variance in SRET change for PSRT vs NTC, $\chi^2(1) = 5.3$, $p = .02$; and indications that there may be significantly less disturbance variance in CES-D change for PSRT vs. NTC, $\chi^2(1) = 3.7$, $p = .054$. Particularly intriguing was the presence of a significantly greater negative self-feedback effect in the NTC group for both CES-D change, $\chi^2(1) = 4.6$, $p = .03$, and especially SRET change, $\chi^2(1) = 18.4$, $p < .001$. This refers to the covariance between successive intervals of change and reflects a tendency for individuals in the NTC to "reverse course" between weeks, e.g., larger initial improvement predicting subsequent setback. The correlation between week 1 and week 2 latent different scores for the CES-D and SRET was substantial and negative for the NTC

group ($\beta = -.34$ for CES-D and $\beta = -.61$ for SRET) and small and positive for the SRET for the PSRT group ($\beta = -.10$ for CES-D and $\beta = .21$ for SRET).

3.5. Analyses in the Intention to Treat Sample

We repeated the primary analyses in the ITT sample, identifying the effect of training on self-schema and on depression symptoms. Where data were missing, only data from the baseline assessment were included. For self-schema, the training group × linear time × valence interaction was not significant, $\beta = 0.11$, 95% CI [−0.07, 0.30], $p = .23$. For depression symptoms, the training group × linear time interaction was not significant, $\beta = 0.42$, 95% CI [−2.08, 2.92], $p = .74$. As in the completers sample, without the interaction the main effect of group remained not significant, $\beta = -1.76$, 95% CI [−4.21, 0.69], $p = .16$, whereas the main effect of time was significant, $\beta = -5.17$, 95% CI [−6.43, −3.92], $p < .001$.

3.6. Subjective benefit and acceptance of training

Because we were piloting a new intervention, we also asked participants who completed the final timepoint of questionnaires to answer several questions about how they liked the training, how easy it was to complete, and how likely they were to recommend the training. These results are included in Table 3. Participants in both conditions generally liked the training (90%), felt that it helped them at least a moderate amount (73%), and felt that it met or exceeded their expectations (91%). Many in the PSRT condition ($N = 22$, 51%), felt that some of the cued scenarios were difficult to imagine. Participants in the NTC condition had marginally less difficulty, with only 14 (33%) reporting that they found imagining scenarios difficult.

4. Discussion

These findings support the hypothesis that positive imagery training would reduce self-referential bias, as measured by the self-referential encoding task (SRET). In two randomly assigned groups of participants, over the course of a two-week period, participants in each group completed the SRET three times and recorded cues every other day according to their training condition. Analyses showed that participants in the active PSRT condition showed a greater increase in the number of positive words they endorsed as self-descriptive, in comparison to the neutral NTC condition. Participants in both conditions showed a decrease in the number of negative words they endorsed as self-descriptive.

These results support the idea that repeated exposure to positive self-referent information, even if imaginal, might alter subsequent self-referent information processing. The idea that self-referentiality influences memory is not new; it is one of the key points underlying investigation of the self-reference effect (Klein et al., 1989; Hamami et al., 2011; Derry & Kuiper, 1981). The self-reference effect argues that information relating to one's own experience can facilitate memory (Klein et al., 1989)—thus, it follows that the repeated exposure in the PSRT condition may provide access to positive self-referential scenes. Importantly, the current work suggests that a focus on self-referential future events may be beneficial to biased self-referenced processing in depression. That is, this link between self-reference and memory appears to be both correlational and causal. Having salient, positive

personal experiences increases the endorsement of positive terms in self-evaluation. It may also improve the salience of positive self-referent imagery and increase the self-referentiality of those positive concepts.

The active training focused on positive imagery, and participants in this condition showed greater improvements in the number of positive words endorsed as self-referent on the SRET. We did not include measures of positive affect, and thus do not know whether these participants would also have increased substantially on such questionnaires. Future work may wish to include such measures. Both trainings reduced endorsement of negative words on the SRET. The NTC training succeeded in training participants away from negative stimuli to a similar degree with the PSRT group. Both groups also showed improvement in positive self-referent processing, although the PSRT group showed a significantly greater increase in endorsement of positive adjectives on the SRET than the NTC group. We are unable to conclude whether the results of the NTC group are due to the general reduction over time of elevated depressive symptoms in a non-clinical sample, to a placebo effect, or to an unintentional efficacy of the neutral condition. While a no-training control group might have permitted a comparison with the ameliorative effects of time alone, in the absence of such a condition, we cannot be certain that the improvements in depression are due to the interventions *per se*.

Contrary to our predictions, the PSRT condition was not associated with a greater reduction in depression symptoms, as measured by the CES-D, than the NTC condition. Both groups saw a reduction in depression symptoms over the course of the study, and this reduction was not significantly different between groups. Given the prediction that there is a causal relationship between self-schema and depression, this lack of differentiation is surprising. It is possible that the changes in self-referential processing were only a small component of self-schema, or that change in depression symptoms would have occurred after a longer follow-up. It is also possible that the NTC condition was so closely matched to the PSRT that the heterogeneous course of depression symptoms made differences between individuals difficult to identify in a small sample. Although we believe that the active ingredients of the PSRT distinguished it from the NTC, it is important to note that both intervention conditions may have, in fact, been effective at reducing depressive symptoms.

Depression change was nonetheless strongly correlated with change on the SRET. Bivariate latent difference scores models, integrating both change in the week from T0 to T1 and change in the week from T1 to T2, demonstrated that a reduction in a negatively-valenced summary score of the SRET predicted a reduction in the CES-D. For both groups, both constructs changed together, with SRET change explaining approximately a quarter of the variance in CES-D change. Although the current sample size was small relative to what is conventionally expected for a structural equation model, the LDS model is a somewhat special case that calculates difference scores based on a regression framework. Most of the paths in the model are pre-specified with fixed coefficients, and power is helped here by the longitudinal, within-subject design and especially the two-groups design. As described in the methods and in greater detail in the supplementary materials, a Monte Carlo simulation demonstrated that our power to detect a significant group difference ($p < .05$) in the causal

path from   SRET to   CESD was 87.9%. Notably, we did not detect a group difference for this path despite adequate power.

Despite the absence of treatment group differences in mean change or the relationship between SRET change and CES-D change, there was evidence that the quality of change may have been different, with the NTC group showing much greater variance in SRET and CES-D change. The starkest difference between groups was a large negative feedback correlation between initial and subsequent change for the NTC group that was minimal (and positive) for the PSRT group. This suggests that the PSRT may have disrupted a negative feedback process governing normal fluctuations in self-schema and depression, without altering the fundamental relationship between self-schema change and depression change. Extrapolating from these dynamics, one speculative hypothesis would be that large observed gains in the NTC group would be lost in the subsequent weeks, whereas the gains in the PSRT group would prove more enduring—indeed, individuals in the PSRT group who showed improvement after the first week were more likely to continue to improve. However, the clinical significance of this finding of reduced variability and disrupted self-feedback—if any—is unknown, and we would caution that this a serendipitous finding from an exploratory analysis. A longer period of observation with a larger sample will be needed to test this idea.

Work in a larger sample could determine whether improvement was moderated by baseline severity of biased self-reference, as has been shown to be the case in larger studies investigating baseline depression severity (Bower et al., 2013). In the current study, the simultaneous change seen in SRET and CES-D supports the theory that self-referent processing plays a role in maintaining depression, given that change on the SRET strongly predicted change in symptomatology as seen in the LDS models. Indeed, such a finding fits with the causal directions predicted by the cognitive model (Beck, 1967), with negatively biased processing (and negative self-schema) helping to maintain depressive biases. Although these findings are preliminary, future work should continue to investigate the causal nature of this relationship. Moreover, we hope that future work may be able to create a granular test for self-reference such that researchers could test the ability of change in self-reference to mediate change in depressive symptoms. Such a study would ideally require a more frequent assessment schedule (e.g., daily measurements of self-schema and mood) than in the current study.

One key element to the PSRT employed in this study was the aim of creating an intervention that participants could tailor to themselves (as per Hertel & Mathews, 2011). This intervention could thus be accessible to a broad range of participants, improving ease of use. Although many participants reported thinking the training was useful, some in the active PSRT condition also reported having difficulty imagining the cues. Such difficulty may be an active ingredient in the efficacy of imagery interventions (Holmes et al., 2016), but is also worth considering carefully in intervention development, given research suggesting that imagination may mediate the success of such interventions (Rohrbacher et al., 2014). Participants' difficulty may be based in the very nature of depression. As such, future work using imagery interventions may provide more guidance to participants, e.g., providing audio scripts—which could improve participants' own imaginings by providing them with

scaffolding on which to build. Some studies have suggested that ambiguity is a necessary component of imagery research, rather than simply asking participants to imagine positive situations (Clarke et al., 2014). The current study's focus on vague cues, instead of specific situations such as falling and breaking a ski rather than a bone (a "positive" interpretation), may result in imaginal work that is similarly ambiguous. However, further work may help in understanding how ambiguity in unspecified cues differs from ambiguity in detailed scenarios.

This study was limited in part by its taking place in a wholly online environment. Participants were not required to have any contact with the experimenters. Although they received emails every other day, these emails were automated, and many participants never emailed the experimenters directly. (The automation also has the effect of resulting in what is, in essence, a double-blind experiment—reducing many potential sources of bias.) Such rare communication limits the ability to generalize the results to in-person therapy, although it is analogous to internet interventions that often have relatively little levels of in-person support (Meyer et al., 2009). Nonetheless, we suggest that the SRET may be very useful in a therapeutic context as an add-on module, and that positive imagery techniques could be useful as an adjunctive treatment alongside CBT or other psychotherapies. Especially given that many participants reported finding even the neutral intervention useful, its appeal in terms of accessibility and ease of use is important to consider.

We also acknowledge that the study had a high degree of participant attrition, including much that occurred after randomization had occurred. Although such attrition is to be expected in online studies, it is unclear whether participants solely did not find the compensation worth their while, or whether other factors beyond those we controlled reduced study completion. This may simply be a peril of online psychology research with such small degrees of participant investment (Postel et al., 2011). However, future work may consider moving randomization to the first follow-up, to limit training to participants who have demonstrated a degree of commitment to the research project; it may also consider increasing methods for retaining participants. In clinical trials attempting to demonstrate treatment efficacy, intention to treat analyses are paramount (Newell, 1992). However, we focus in this paper on a completers analysis because the primary focus of this project was mechanistic.

In short, we believe that the current study provides further support for the role of negative self-referent processing in maintaining depression. It additionally suggests that positive imagery may be a promising approach to increase positive self-referent processing. Continuing to explore how positive imagery may increase positive self-schema—and reduce negative self-schema—is likely important for reducing depressive symptoms. These results may both help us in understanding how the symptoms of depression may be impacted by the salience of emotional information, and in understanding ways of modifying biased views of the self. Modifying such views may be vital to developing treatments targeting depressed mood.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Auerbach RP, Stanton CH, Proudfit GH, & Pizzagalli DA (2015). Self-referential processing in depressed adolescents: A high-density event-related potential study. Journal of Abnormal Psychology, 124, 233–45. doi:10.1037/abn0000023. [PubMed: 25643205]

Bates D, Mächler M, Bolker B, & Walker S (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67, 1–48. doi:10.18637/jss.v067.i01.

Beck AT (1967). Depression: Clinical, experimental, and theoretical aspects. University of Pennsylvania Press.

Beck AT (1979). Cognitive therapy and the emotional disorders. Penguin Books.

Boker SM, Neale MC, Maes HH, Wilde MJ, Spiegel M, Brick TR, Estabrook R, Bates TC, Mehta P, von Oertzen T, Gore RJ, Hunter MD, Hackett DC, Karch J, Brandmaier AM, Pritikin JN, Zahery M, Kirkpatrick RM, Wang Y, Driver C, Massachusetts Institute of Technology, Johnson SG, Association for Computing Machinery, Kraft D, Wilhelm S, & Manjunath BG (2017). OpenMx 2.7.10 User Guide.

Boland J, Riggs KJ, & Anderson RJ (2017). A brighter future: The effect of positive episodic simulation on future predictions in non-depressed, moderately dysphoric & highly dysphoric individuals. Behaviour Research and Therapy, (p. Advance online publication). doi:10.1016/j.brat. 2017.10.010.

Bower P, Kontopantelis E, Sutton A, Kendrick T, Richards DA, Gilbody S, Knowles S, Cuijpers P, Andersson G, Christensen H, Meyer B, Huibers M, Smit F, van Straten A, Warmerdam L, Barkham M, Bilich L, Lovell K, & Liu ET-H (2013). Influence of initial severity of depression on effectiveness of low intensity interventions: Meta-analysis of individual patient data. BMJ, 346, f540. doi:10.1136/bmj.f540. [PubMed: 23444423]

Bradley BP, Mogg K, & Williams R (1995). Implicit and explicit memory for emotion-congruent information in clinical depression and anxiety. Behaviour Research and Therapy, 33, 755–770. doi: 10.1016/0005-7967(95)00029-W. [PubMed: 7677713]

Bradley MM, & Lang PJ (2010). Affective norms for English words (ANEW): Stimuli, instruction manual, and affective ratings (Technical report C-2). Technical Report The Center for Research in Psychophysiology, University of Florida Gainesville, FL.

Christensen H, Griffiths KM, & Farrer L (2009). Adherence in internet interventions for anxiety and depression: Systematic review. Journal of Medical Internet Research, 11, e13. doi:10.2196/jmir. 1194. [PubMed: 19403466]

Clarke PJ, Nanthakumar S, Notebaert L, Holmes EA, Blackwell SE, & MacLeod C (2014). Simply imagining sunshine, lollipops and rainbows will not budge the bias: The role of ambiguity in interpretive bias modification. Cognitive Therapy and Research, 38, 120–131. doi:10.1007/s10608-013-9564-x. [PubMed: 24634553]

Connolly SL, Abramson LY, & Alloy LB (2015). Information processing biases concurrently and prospectively predict depressive symptoms in adolescents: Evidence from a self-referent encoding task. Cognition and Emotion, (pp. 1–11). doi:10.1080/02699931.2015.1010488.

Dainer-Best J (2017). Guided imagery training and self-referential processing. URL: https://clinicaltrials.gov/ct2/show/NCT03056963identifierNCT03056963.

Dainer-Best J, Lee HY, Shumake JD, Yeager DS, & Beevers CG (2018). Determining optimal parameters of the self referent encoding task: A large-scale examination of self-referent cognition and depression. Psychological Assessment,. doi:10.1037/pas0000602.

Dainer-Best J, Trujillo LT, Schnyer DM, & Beevers CG (2017). Sustained engagement of attention is associated with increased negative self-referent processing in major depressive disorder. Biological Psychology, 129, 231–241. doi:10.1016/j.biopsycho.2017.09.005. [PubMed: 28893596]

Derry PA, & Kuiper NA (1981). Schematic processing and self-reference in clinical depression. Journal of Abnormal Psychology, 90, 286–297. doi:10.1037/0021-843X.90.4.286. [PubMed: 7264058]

Disner SG, Shumake JD, & Beevers CG (2016). Self-referential schemas and attentional bias predict severity and naturalistic course of depression symptoms. Cognition and Emotion, (pp. 1–13). [PubMed: 26513687]

Doost HTN, Moradi AR, Taghavi MR, Yule W, & Dalgleish T (1999). The development of a corpus of emotional words produced by children and adolescents. Personality and Individual Differences, 27, 433–451. doi:10.1016/S0191-8869(98)00253-0.

Dozois DJ, Bieling PJ, Patelis-Siotis I, Hoar L, Chudzik S, McCabe K, & Westra HA (2009). Changes in self-schema structure in cognitive therapy for major depressive disorder: A randomized clinical trial. Journal of Consulting and Clinical Psychology, 77, 1078–1088. doi:10.1037/a0016886. [PubMed: 19968384]

Ferrari GR, Möbius M, van Opdorp A, Becker ES, & Rinck M (2016). Can't look away: An eye-tracking based attentional disengagement training for depression. Cognitive Therapy and Research, 40, 672–686. doi:10.1007/s10608-016-9766-0. [PubMed: 27688520]

Fox J (2003). Effect displays in R for generalised linear models. Journal of Statistical Software, 8, 1–27. doi:10.18637/jss.v008.i15.

Goldin PR, Jazaieri H, Ziv M, Kraemer H, Heimberg RG, & Gross JJ (2013). Changes in positive self-views mediate the effect of cognitive-behavioral therapy for social anxiety disorder. Clinical Psychological Science, 1, 301–310. doi:10.1177/2167702613476867. [PubMed: 25541599]

Gotlib IH, & Joormann J (2010). Cognition and depression: Current status and future directions. Annual Review of Clinical Psychology, 6, 285–312. doi:10.1146/annurev.clinpsy.121208.131305.

Gotlib IH, & Krasnoperova E (1998). Biased information processing as a vulnerability factor for depression. Behavior Therapy, 29, 603–617. doi:10.1016/S0005-7894(98)80020-8.

Gupta R, & Kar BR (2012). Attention and memory biases as stable abnormalities among currently depressed and currently remitted individuals with unipolar depression. Frontiers in Psychiatry, 3, 99. doi:10.3389/fpsyt.2012.00099. [PubMed: 23189063]

Halekoh U, & Højsgaard S (2014). A kenward-roger approximation and parametric bootstrap methods for tests in linear mixed models – the R package pbkrtest. Journal of Statistical Software, 59, 1–30. doi:10.18637/jss.v059.i09. [PubMed: 26917999]

Hall KAA, Raedt RD, Timpano KR, & Joormann J (2018). Positive memory enhancement training for individuals with major depressive disorder. Cognitive Behaviour Therapy, 47, 155–168. doi:10.1080/16506073.2017.1364291. [PubMed: 28826327]

Hallion LS, & Ruscio AM (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. Psychological Bulletin, 137, 940–958. doi:10.1037/a0024355. [PubMed: 21728399]

Hamami A, Serbun SJ, & Gutchess AH (2011). Self-referencing enhances memory specificity with age. Psychology and Aging, 26, 636–646. doi:10.1037/a0022626. [PubMed: 21480719]

Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, & Conde JG (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. Journal of Biomedical Informatics, 42, 377–381. doi:10.1016/j.jbi.2008.08.010. [PubMed: 18929686]

Hertel PT, & Mathews A (2011). Cognitive bias modification past perspectives, current findings, and future applications. Perspectives on Psychological Science, 6, 521–536. doi:10.1177/1745691611421205. [PubMed: 26168375]

Hitchcock C, Mueller V, Hammond E, Rees C, Werner-Seidler A, & Dalgleish T (2016). The effects of autobiographical memory flexibility (Mem-Flex) training: An uncontrolled trial in individuals in

remission from depression. Journal of Behavior Therapy and Experimental Psychiatry, 52, 92–98. doi:10.1016/j.jbtep.2016.03.012. [PubMed: 27058165]

Hollis S, & Campbell F (1999). What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ, 319, 670–674. doi:doi.org/10.1136/bmj.319.7211.670. [PubMed: 10480822]

Holmes EA, Blackwell SE, Burnett Heyes S, Renner F, & Raes F (2016). Mental imagery in depression: Phenomenology, potential mechanisms, and treatment implications. Annual Review of Clinical Psychology, 12, 249–280. doi:10.1146/annurev-clinpsy-021815-092925.

Holmes EA, Lang TJ, & Shah DM (2009). Developing interpretation bias modification as a "cognitive vaccine" for depressed mood: Imagining positive events makes you feel better than thinking about them verbally. Journal of Abnormal Psychology, 118, 76–88. doi:10.1037/a0012590. [PubMed: 19222316]

Holmes EA, & Mathews A (2010). Mental imagery in emotion and emotional disorders. Clinical Psychology Review, 30, 349–362. doi:10.1016/j.cpr.2010.01.001. [PubMed: 20116915]

Joormann J, Waugh CE, & Gotlib IH (2015). Cognitive bias modification for interpretation in major depression: Effects on memory and stress reactivity. Clinical Psychological Science, 3, 126–139. doi:10.1177/2167702614560748. [PubMed: 25593790]

Klein SB (2013). The complex act of projecting oneself into the future. Wiley Interdisciplinary Reviews: Cognitive Science, 4, 63–79. doi:10.1002/wcs.1210. [PubMed: 26304175]

Klein SB, Loftus J, & Burton HA (1989). Two self-reference effects: The importance of distinguishing between self-descriptiveness judgments and autobiographical retrieval in self-referent encoding. Journal of Personality and Social Psychology, 56, 853–865. doi:10.1037/0022-3514.56.6.853.

LeMoult J, Colich N, Joormann J, Singh MK, Eggleston C, & Gotlib IH (2017a). Interpretation bias training in depressed adolescents: Near- and far-transfer effects. Journal of Abnormal Child Psychology, (pp. 1–9). doi:10.1007/s10802-017-0285-6. [PubMed: 27113216]

LeMoult J, Kircanski K, Prasad G, & Gotlib IH (2017b). Negative self-referential processing predicts the recurrence of major depressive episodes. Clinical Psychological Science, 5, 174–181. doi: 10.1177/2167702616654898. [PubMed: 28286705]

Loeys T, Moerkerke B, & Vansteelandt S (2015). A cautionary note on the power of the test for the indirect effect in mediation analysis. Frontiers in Psychology, 5, 1549. doi:10.3389/fpsyg. 2014.01549. [PubMed: 25628585]

MacLeod C, Koster EH, & Fox E (2009). Whither cognitive bias modification research? Commentary on the special section articles. Journal of Abnormal Psychology, 118, 89–99. doi:10.1037/ a0014878. [PubMed: 19222317]

Meyer B, Berger T, Caspar F, Beevers CG, Andersson G, & Weiss M (2009). Effectiveness of a novel integrative online treatment for depression (Deprexis): Randomized controlled trial. Journal of Medical Internet Research, 11, e15. doi:10.2196/jmir.1151. [PubMed: 19632969]

Moher D, Schulz KF, & Altman DG (2001). The CONSORT statement:Revised recommendations for improving the quality of reports of parallel-group randomised trials. The Lancet, 357, 1191–1194. doi:10.1016/S0140-6736(00)04337-3.

Neale MC, Hunter MD, Pritikin JN, Zahery M, Brick TR, Kickpatrick RM, Estabrook R, Bates TC, Maes HH, & Boker SM (2016). OpenMx 2.0: Extended structural equation and statistical modeling. Psychometrika, 81, 535–549. doi:10.1007/s11336-014-9435-8. [PubMed: 25622929]

Neshat-Doost HT, Dalgleish T, Yule W, Kalantari M, Ahmadi SJ, Dyregrov A, & Jobson L (2013). Enhancing autobiographical memory specificity through cognitive training: An intervention for depression translated from basic science. Clinical Psychological Science, 1, 84–92. doi: 10.1177/2167702612454613.

Newell DJ (1992). Intention-to-treat analysis: Implications for quantitative and qualitative research. International Journal of Epidemiology, 21, 837–841. doi:10.1093/ije/21.5.837. [PubMed: 1468842]

Newsom JT (2015). Longitudinal structural equation modeling: A comprehensive introduction chapter Latent Difference Score Models. (p. 258). New York, NY: Routledge.

Phillips WJ, Hine DW, & Thorsteinsson EB (2010). Implicit cognition and depression: A meta-analysis. Clinical Psychology Review, 30, 691–709. doi:10.1016/j.cpr.2010.05.002. [PubMed: 20538393]

Postel MG, de Haan HA, Ter Huurne ED, van der Palen J, Becker ES, & de Jong CA (2011). Attrition in web-based treatment for problem drinkers. Journal of Medical Internet Research, 13, e117. doi:10.2196/jmir.1811. [PubMed: 22201703]

Pritikin JN, Hunter MD, & Boker SM (2015). Modular open-source software for Item Factor Analysis. Educational and Psychological Measurement, 75, 458–474. [PubMed: 27065479]

Quilty LC, Dozois DJ, Lobo DS, Ravindran LN, & Bagby RM (2014). Cognitive structure and processing during cognitive behavioral therapy vs. pharmacotherapy for depression. International Journal of Cognitive Therapy, 7, 235–250. doi:10.1521/ijct.2014.7.3.235.

Radloff LS (1977). The CES-D scale: A self-report depression scale for research in the general population. Applied Psychological Measurement, 1, 385–401. doi:10.1177/014662167700100306.

Raes F, Williams JMG, & Hermans D (2009). Reducing cognitive vulnerability to depression: A preliminary investigation of MEmory Specificity Training (MEST) in inpatients with depressive symptomatology. Journal of Behavior Therapy and Experimental Psychiatry, 40, 24–38. doi:10.1016/j.jbtep.2008.03.001. [PubMed: 18407245]

Rohrbacher H, Blackwell SE, Holmes EA, & Reinecke A (2014). Optimizing the ingredients for imagery-based interpretation bias modification for depressed mood: Is self-generation more effective than imagination alone? Journal of Affective Disorders, 152–154, 212–218. doi:10.1016/j.jad.2013.09.013.

Santor DA, Zuroff DC, Ramsay J, Cervantes P, & Palacios J (1995). Examining scale discriminability in the BDI and CES-D as a function of depressive severity. Psychological Assessment, 7, 131–139. doi:10.1037/1040-3590.7.2.131.

Thurston MD, Goldin P, Heimberg R, & Gross JJ (2017). Self-views in social anxiety disorder: The impact of CBT versus MBSR. Journal of Anxiety Disorders, 47, 83–90. doi:10.1016/j.janxdis.2017.01.001. [PubMed: 28108059]

Watkins ER, Baeyens CB, & Read R (2009). Concreteness training reduces dysphoria: Proof-of-principle for repeated cognitive bias modification in depression. Journal of Abnormal Psychology, 118, 55–64. doi:10.1037/a0013642. [PubMed: 19222314]

Wells TT, & Beevers CG (2010). Biased attention and dysphoria: Manipulating selective attention reduces subsequent depressive symptoms. Cognition and Emotion, 24, 719–728. doi:10.1080/02699930802652388.

Wickham H (2009). ggplot2: Elegant Graphics for Data Analysis Springer-Verlag New York. URL: http://ggplot2.org.

Wickham H (2017). tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions. URL: https://CRAN.R-project.org/package=tidyrrpackage version 0.6.1.

Wickham H, Francois R, Henry L, & Müller K (2015). dplyr: A Grammar of Data Manipulation. URL: https://CRAN.R-project.org/package=dplyrrpackage version 0.7.1.

Highlights

- People with depressed mood describe themselves more negatively and less positively

- A positive imagery training increased positive self-reference over two weeks

- The positive training also focused on the self, the future, and specific imagery

- Compared to a neutral control, the positive imagery was more effective at improving positive self-referent processing

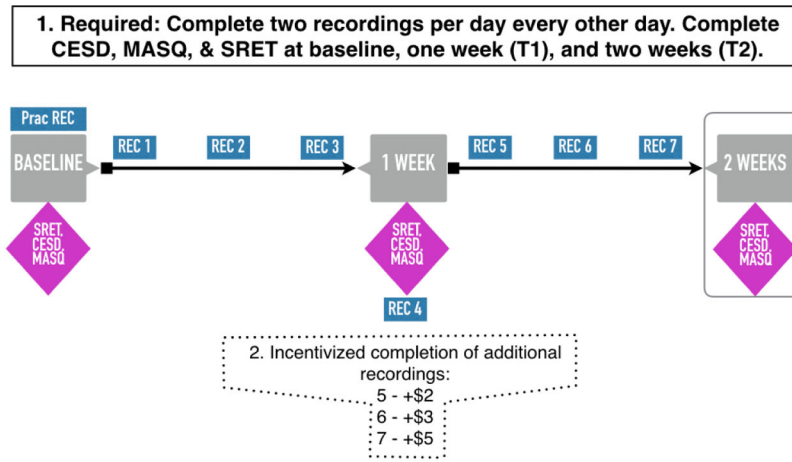- Both conditions showed less negative self-reference and depressive symptoms

Figure 1:
Assessment schedule showing how SRET, CES-D, and recordings were administered.
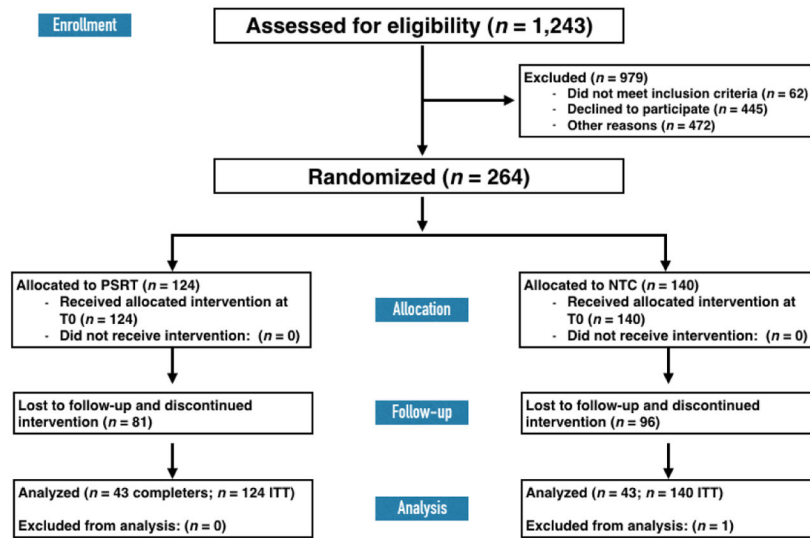
Figure 2:
CONSORT flow chart. There were 1,243 visitors who began the eligibility screen; their uniqueness cannot be assessed. Following randomization, the interventions were accessed 264 unique times. All such participants were included in either ITT (intention-to-treat) or completer samples.
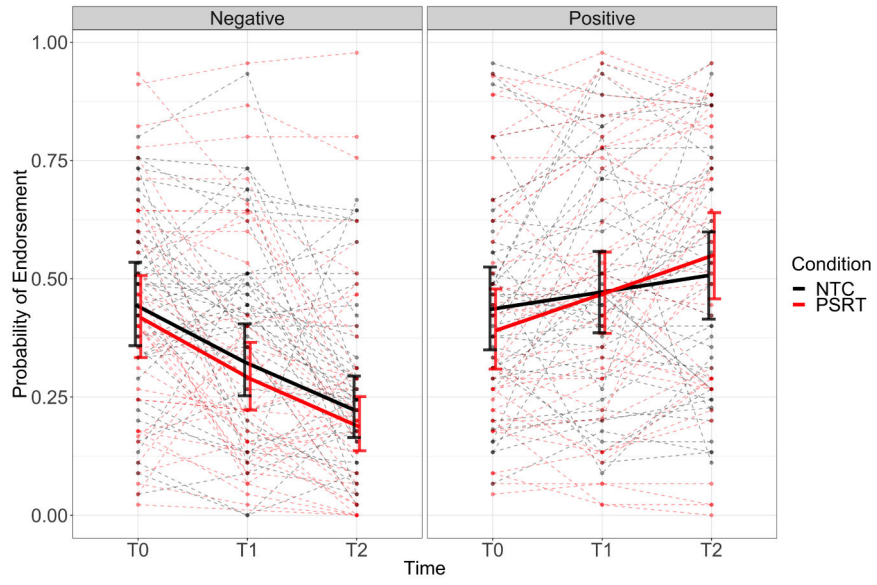
Figure 3:
Change in probability of endorsing words over time, presented separately for negative (left) and positive (right) words. The PSRT training condition is presented in red; the NTC training condition in black. Dashed lines show individual participants' trajectories. Solid, thick lines show the linear trend of time as estimated by the logistic regression mixed effects model, with error bars demonstrating 95% CIs around the estimates at each time-point. The probability of endorsing negative words fell for each group, while the probability of endorsing positive words increased more for the PSRT group than for the NTC group.
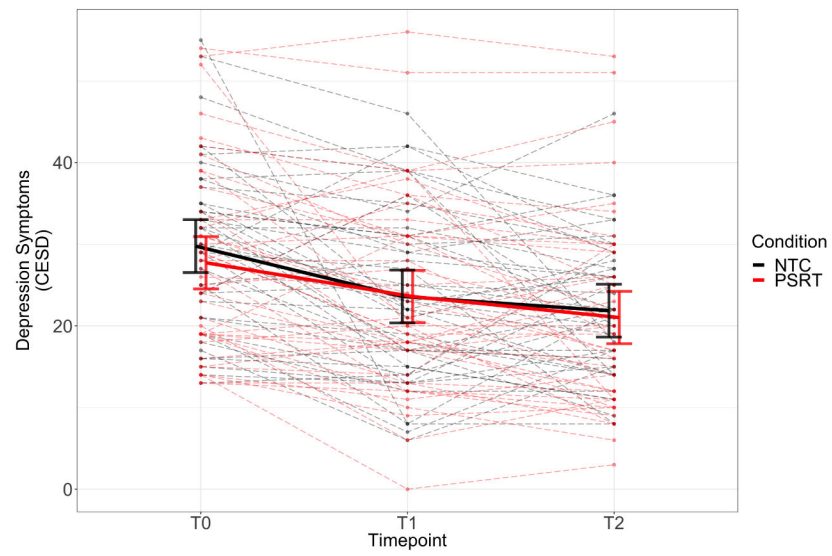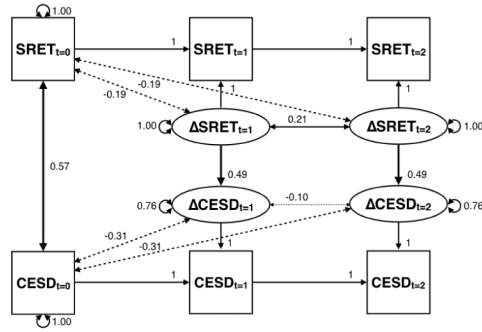
Figure 4:
Change in depression symptoms over time. The PSRT group is in red; the NTC in black. Dashed lines show individual participants' CES-D trajectories. Solid, thick lines show the linear trend in group means over time as estimated by the regression model, with error bars demonstrating 95% CIs around the estimates at each time-point.
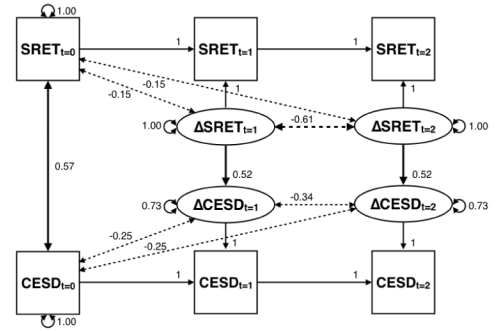
Figure 5:

Structural equation models (SEM) created separately for PSRT and NTC groups. Lines with double-headed arrows represent covariance between variables, and lines with single-headed arrows represent causal paths. Solid lines indicate positive relationships, and dashed lines indicate negative relationships. Thicker lines represent stronger relationships. Means of SRET and CES-D at T0 and the latent change variables (ΔSRET and ΔCES-D) were also modeled but not shown in diagram (see Table 2). Mean depression, as measured by the CES-D, decreased over time for both groups. Mean SRET (measured as negative – positive) also decreased over time for both groups. Primary differences between groups' LDS models are noted in the relationship between change in SRET from T1 to T2, and change in CES-D from T1 to T2. In the PSRT group, change in SRET and CES-D at T1 is relatively decoupled from change in SRET and CES-D at T2; in the NTC group, change in SRET and CES-D at T1 is significantly and negatively correlated with change in SRET and CES-D at T2. This indicates, for example, that individuals in the NTC group who show more initial improvement are predicted to subsequently worsen (a negative feedback effect), whereas the effect was, if anything, reversed in the PSRT group. Lines in Figures 3 and 4 demonstrate varying individual trajectories of participants' change.

Table 1:

Characteristics and symptom profiles of completer participants at baseline. CES-D = Center for Epidemiologic Studies — Depression Scale, which ranges from 0 to 60. SRET = Self-Referential Encoding Task.

	PSRT (*N* = 44)	NTC (*N* = 43)
Age, mean (SD)	25.5 (6.2)	27.4 (7.6)
Gender		
Female	35 (80%)	37 (86%)
Male	8 (18%)	6 (14%)
Genderqueer	1 (2%)	0
Race		
White	29 (65.9%)	28 (65.1%)
Black	5 (11.3%)	3 (7.0%)
American Indian	0	1 (2.3%)
Asian	5 (11.3%)	7 (16.3%)
No response	5 (11.3%)	4 (9.3%)
Hispanic/Latino	4 (9%)	5 (12%)
CES-D, mean (SD)	27.7 (11.3)	29.8 (10.2)
Current self-reported Major Depressive Episode	14 (31.8%)	12 (28%)
SRET, mean (SD) per parameter		
# Positive Words Endorsed	20.5 (11.7)	19.9 (10.3)
# Negative Words Endorsed	19.3 (10.6)	21.5 (9.7)
RT to Positive Words (ms)	1403 (1064)	1260 (559)
RT to Negative Words (ms)	1479 (1954)	1225 (603)

Table 2:

For the latent difference score (LDS) models, this table presents the standardized (β) and unstandardized (b) coefficients, with their standard errors (SE). Standardized values are standardized against the variance.[**] The third column presents the χ^2 and p values from the likelihood ratio test comparing models in which that parameter was either fixed to equality between groups or free to vary between groups. Coefficients that were constrained to equality by the final model are highlighted in gray and are, by definition, equivalent between conditions.

	PSRT		NTC			
	b (*SE*)	β	*b* (*SE*)	β	$\chi^2(1)$	*p*
Path						
SRET change → CES-D change	0.35 (0.05)	0.49	0.35 (0.05)	0.52	0.1	.73
Covariances						
SRET ↔ CES-D at baseline	75 (15)	0.57	75 (15)	0.57	1.3	.26
Regress-to-mean SRET	−27.5 (9.2)	−0.19	−27.5 (9.2)	−0.15	0.6	.45
Regress-to-mean CES-D	−22.5 (4.6)	−0.31	−22.5 (4.6)	−0.25	0.0	.94
Self-feedback SRET change	22.8 (17)	0.21	−117 (36)	−0.61	18.4	< .001
Self-feedback CES-D change	−5.91 (6.5)	−0.10	−30.9 (12)	−0.34	4.6	.033
Means						
SRET at baseline	39.9 (1.5)		39.9 (1.5)		0.2	.68
CES-D at baseline	28.5 (1.1)		28.5 (1.1)		0.4	.54
SRET change/week	−5.8 (0.79)		−5.8 (0.79)		0.0	.92
CES-D change/week	−1.82 (0.55)		−1.82 (0.55)		0.0	.98
Variances						
SRET at baseline	179 (27)	1.0[*]	179 (27)	1.0[*]	0.0	.85
CES-D at baseline	90.1 (12)	1.0[*]	90.1 (12)	1.0[*]	0.2	.62
SRET change/week	111 (17)	1.0[*]	192 (36)	1.0[*]	5.3	.022
CES-D change/week	43.2 (6.5)	0.76	66 (12)	0.73	3.7	.054

[*] The coefficients are standardized against the variance and are by definition 1 for all exogenous variables.

[**] Because the variances were not always constrained to be equal, an identical unstandardized path coefficient may have different standardized coefficients because the reference for that standardization is different.

Table 3:

Subjective benefit and acceptance of training. Responses of completers participants to the end-of-task questionnaire on their experience of the treatment.

	PSRT	NTC
Overall impression of the program		
Positive (4-6)	37	40
Neutral (3)	6	2
Negative (1-2)	0	1
Feeling that the program helped		
Helped a lot	6	6
Helped a little	28	23
Did not help	9	14
Meeting expectations		
Exceeded	6	8
Met	33	31
Failed to meet	4	4
Likelihood of recommending		
Yes	31	30
No	12	12
Ease in imagining cued scenarios		
Easy (4-5)	19	24
Neutral (3)	1	5
Difficult (1-2)	22	14
Understanding of training		
Straightforward	32	33
Confusing	15	11
Ease of technology interface		
Easy (4-5)	41	39
Neutral (3)	0	0
Difficult (1-2)	2	4
Technical Problems		
None	33	29
Yes	6	10
Yes, and it got in the way	4	4