Shifting Negative Prospection With Online Cognitive Bias Modification: A Randomized Controlled Trial

Jeremy W. Eberle

Charlottesville, Virginia

A predissertation research project presented to the graduate faculty of the University of Virginia in candidacy for the degree of Master of Arts

Department of Psychology

University of Virginia

Abstract

Objective: Emotional disorders often include negative simulations of the future, termed *negative prospection*. The present study tested an online cognitive bias modification program designed to train more positive future thinking in community participants. ***Method:*** 958 adults (73.3% female, 86.5% White, 83.4% from United States), most (74.1%) with a likely anxiety or depressive disorder, were randomized to one of five conditions: two positive conditions with emotionally ambiguous future scenarios that ended positively 90% of the time after first either negating a negative outcome (n = 147) or not (n = 177), two 50/50 conditions that ended positively (50% of the time) or negatively (50% of the time) in either blocked (n = 146) or random (n = 173) order, and a control condition with emotionally neutral scenarios (n = 315). Outcomes were assessed at baseline, after each of four training sessions, and at 1-month follow-up. ***Results:*** As hypothesized (preregistration: osf.io/jrst6), participants in positive training improved in negative and positive expectancy bias, self-efficacy, and optimism more than control participants, *d*s and 97.5% CIs = -0.57 [-0.87, -0.27], 0.79 [0.42, 1.15], 0.28 [0.02, 0.53], 0.28 [0.04, 0.51], and, for expectancy bias, more than 50/50 participants. Unexpectedly, participants across all conditions improved comparably in anxiety and depression symptoms and growth mindset. Additionally, no superiority emerged between the control and 50/50 conditions, between the two 50/50 conditions, or between the two positive conditions. ***Conclusions:*** Targeting transdiagnostic negative prospection with a scalable program may improve bias and outlook; however, further validation of outcome measures is required.

Keywords: cognitive bias modification, prospection, expectancy bias, anxiety, depression

Shifting Negative Prospection With Online Cognitive Bias Modification:

A Randomized Controlled Trial

Emotional disorders are prevalent but vastly undertreated. Anxiety disorders and major depression, for example, occur in about 7.3% and 4.7% of people, respectively, at a given time (Baxter, Scott, Vos, & Whiteford, 2013; Ferrari et al., 2013), but only 5-20% of anxious individuals and 7-28% of depressed individuals are receiving treatment (Chisholm et al., 2016). Further, those who do eventually talk with a professional about an anxiety or mood disorder do so after a median delay of 3-30 years and 1-14 years, respectively (Wang et al., 2007). Reducing this treatment gap will require new models for delivering treatments; the dominant individual therapy model is not scalable (Kazdin, 2017).

In addition to developing delivery models that expand the reach of existing treatments, developing models that target mechanisms of change directly—focusing more on procedures for processes than on protocols for syndromes (Hofmann & Hayes, 2019)—may improve efficacy and scalability (Kazdin, 2007). For example, online cognitive bias modification (CBM) programs that target disorder-relevant processing biases can reduce symptoms and, because they typically require no professional contact, can be completed anywhere with Internet access, making them accessible to people who otherwise may face a host of barriers to treatment (Teachman, 2014).

Negative Prospection in Emotional Disorders

One potential transdiagnostic mechanism to target in emotional disorders is *prospection*, or simulations of future events (Gilbert & Wilson, 2007), which may be negative in ways that maintain anxiety and mood disorders (Seligman, Railton, Baumeister, & Sripada, 2013; Roepke & Seligman, 2016). Early cognitive models of anxiety and depression, for example, posited the role of maladaptive prospection in these disorders (Beck, Brown, Steer, Eidelson, & Riskind, 1987). Research has shown that, in contrast to an optimism bias among healthy adults, adults with anxiety and depression tend to expect more negative events in the future and that adults with depression also expect fewer positive events (Miloyan, Pachana, & Suddendorf, 2014).

Whereas negative prospection appears present across the emotional disorders, positive prospection has benefits such as improved emotion regulation (Pham & Taylor, 1999) and problem-solving abilities (e.g., Miloyan & Suddendorf, 2015). A psychotherapist might target rigidly negative future thinking by having a client generate alternative future possibilities, thereby promoting a more positive and flexible prospective style. The present study tested the effectiveness and feasibility of an online interpretation bias training program adapted to target negative future thinking directly (Namaky, Glenn, Eberle, & Teachman, 2019) and delivered to adults around the world using our team's public research website called MindTrails.

Approaches to Manipulating Prospection

Prospection has been manipulated in several ways in prior research. Participants have simulated positive future states by writing about their best possible future selves (Malouff & Shutte, 2017; Meevissen, Peters, & Alberts, 2011) or ordinary positive future events (Quoidbach, Wood, & Hansenne, 2009). In addition, participants have simulated positive future events in response to cue words that were self-generated (Szpunar & Schacter, 2013) or provided (Boland, Riggs, & Anderson, 2018). Participants have also been instructed to imagine positive outcomes or the process of achieving them (Taylor, Pham, Rivkin, & Armor, 1998). Finally, in variations on the ambiguous scenarios paradigm (Mathews & Mackintosh, 2000), a form of interpretation bias training (CBM-I; Jones & Sharpe, 2017), participants have been instructed to imagine being in emotionally ambiguous scenarios that ultimately resolve with a positive or negative ending.

In one variation of the ambiguous scenarios paradigm, positive imagery CBM-I, participants resolve ambiguous present-tense scenarios with positive endings, and some authors have proposed that resolving this ambiguity involves the generation of future imagery (Lang, Blackwell, Harmer, Davison, & Holmes, 2012; Murphy et al., 2017), a form of episodic simulation, which is one mode of prospection (Szpunar, Spreng, & Schacter, 2014). These scenarios, and ambiguous pictures that resolve with a positive caption, can increase the vividness of positive future imagery (Blackwell et al., 2015), behavioral activation (Renner, Ji, Pictet, Holmes, & Blackwell, 2017), and optimism (Murphy et al., 2015).

Targeting Negative Prospection Directly

In another variation of CBM-I using the ambiguous scenarios paradigm, Namaky et al. (2019) developed a web-based program designed to target negative prospection directly with scenarios that described short- and long-term future events, used future tense in resolving the ambiguity of the scenarios, and solicited future predictions in post-scenario comprehension questions. They found that, in college students with more negative future thinking than most of the students screened, participants assigned to the positive and 50/50 (half-positive, half-negative) conditions showed more positive expectancy bias and greater self-efficacy and growth mindset than those in an active control condition. Participants across all conditions improved in anxiety and depression symptoms and in optimism.

Although the initial findings of Namaky et al. (2019) were promising, they were limited to a small college sample; based on two time points for most outcomes, from baseline to a short (1-week) follow-up; and did not show condition differences over time for some outcomes (other CBM-I studies have also found mixed results: Jones & Sharpe, 2017; Menne-Lothmann et al., 2014). Moreover, Namaky et al. did not assess the effectiveness or feasibility of implementing the intervention on a public platform accessible to adults around the world. Evaluating feasibility is critical because recruitment and retention can be difficult for public online interventions. In particular, attrition is common (Eysenbach, 2005); only 10% of users complete a second module at MoodGym, a popular cognitive behavior therapy site, for example (Batterham, Neil, Bennett, Griffiths, & Christensen, 2008). The present study builds on Namaky et al.'s findings by testing a similar program in a larger, broader sample and including mid- and post-treatment assessments and a slightly longer follow-up period (1 month) on a platform easily disseminable to the public.

Overview of Present Study and Hypotheses

The present study, conducted on the MindTrails research website, is a randomized controlled trial of a cognitive bias modification intervention for reducing negative expectancy bias and increasing positive expectancy bias in community adults with negative prospection. Participants were randomized to one of five conditions. Positive Prospection was designed to train more positive future thinking via repeated practice envisioning positive outcomes to emotionally ambiguous, self-relevant future situations. A second positive condition, Positive Prospection + Negation, supplements the envisioning of positive outcomes with a negation of negative outcomes. Two 50/50 conditions, designed to provide practice envisioning different outcomes to emotionally ambiguous future situations without training a contingency, present equal proportions of positive and negative outcomes: 50/50 Blocked uses blocks of alternating valence, whereas 50/50 Random uses a random order. Neutral Control controls for the CBM format of the other conditions but uses situations that lack emotional ambiguity about the future.

We preregistered several directional hypotheses (osf.io/jrst6): First, during treatment participants in the two positive conditions will decrease in negative expectancy bias and anxiety and depression symptoms and increase in positive expectancy bias, self-efficacy, growth mindset, and optimism significantly more than those in Neutral Control. Second, participants in the 50/50 conditions will improve significantly more than those in Neutral Control but less than those in the positive conditions. Third, participants in 50/50 Blocked will improve significantly more than those in 50/50 Random given that the blocked design may promote more flexibility in requiring that they to shift their future thinking in a given block after developing a consistently positive or negative expectation in the prior block. Finally, we preregistered a nondirectional test of improvements between the two positive conditions, expecting that negating negative outcomes could either disconfirm negative expectations and improve efficacy (Seligman et al., 2013) or reinforce negative associations and reduce efficacy (Ouimet et al., 2009).

Method

Participants and Design

Data collection began on May 3, 2017. We analyzed data collected through September 9, 2018, for participants who enrolled on or before March 27, 2018. 4,751 community participants self-selected to complete a screening for an online study "to encourage healthier thinking about the future for people who tend to expect things will not turn out well" on the MindTrails Project website (https://mindtrails.virginia.edu), suitable for computers, tablets, and smartphones. 1,221 participants at least 18 years of age and with index scores more than 0.5 standard deviations below the mean score on the Expectancy Bias Task in a prior sample (see below) were eligible, provided informed consent, created an account, and were randomized to one of five conditions (see Section S1.1 in supplement[1] for details). (Procedures for all participants, who were not told their condition, were identical from the point of randomization to the start of the first session.) Participants were encouraged to complete a pretreatment assessment within 2 days. After they did, they were immediately given the opportunity to begin the first treatment session. 971 began

[1]Sections S1-S3, Tables S1-S10, and Figures S1-S7 are available in the online supplement.

the first session (i.e., viewed at least the first training scenario); 13 of these (1 who had submitted a blank screening and 12 who had repeated the screening until they became eligible) were excluded, forming an intent-to-treat (ITT) sample of 958 participants. 289 of these completed all four treatment sessions and form the per-protocol (PP) sample. See Figure 1 for the participant flow. The University of Virginia Institutional Review Board approved all research procedures.

ITT participants were primarily female (73.3%), White (86.5%), Not Hispanic or Latino (88.3%) adults (M = 40.94 years, SD = 13.41)[2] from the United States (83.4%) who had finished at least some college (93.1%) or at least some graduate school (46.3%). Most were working full (53.3%) or part time (11.6%) or were students (11.3%). Half annually earned less than $50,000 (28.6%) or between $50,000 and $100,000 (24.6%); a third earned more than $100,000 (31.7%). Most were in a relationship (63.7%) or single (23.2%). Most (74.1%) scored above thresholds for a likely anxiety disorder (62.9%), depressive disorder (50.0%), or both (38.8%; Patient Health Questionnaire-4, see below). See Table S2 for full demographic information.

Recruitment. The MindTrails website and present study were advertised through online, university press releases; local and statewide news (e.g., radio interviews); Craigslist postings; emails to clinicians; flyers at university counseling centers; and the ongoing MindTrails Project Facebook page. A link to the original MindTrails study (Ji et al., 2019) was posted on ClinicalTrials.gov (NCT02382003).

Treatment and assessment schedule. Participants were asked to complete four training sessions (two per week, 2-4 days apart). Assessments were given in a fixed order immediately after each session and at 1-month follow-up (because training and assessment occurred in an

[2]Two participants reported birth years of 1900 and 2017 and are excluded from this mean and standard deviation. Three participants reported birth years suggesting that they were less than 18 years of age; the participants are included in this mean and standard deviation. All five of these participants are included in subsequent analyses because before enrolling in the study they had checked a box confirming they were at least 18 years of age.

established sequence, non-PP participants were also lost to follow-up). Participants had to wait 2 days before starting the next training session and 30 days before starting the follow-up assessment; they could then start the next component at any time. Participants had the option of receiving an email or text reminder when the component was due, and if they completed only part of a component, they continued it the next time they returned (see Sections S1.2 and S1.3).

Outcome Measures

Expectancy bias. Expectancy bias was assessed with a modified Expectancy Bias Task (Namaky et al., 2019), a reading judgment task that assesses tendencies to expect positive or negative events. Participants read and imagined themselves in four scenarios, each containing a title, orienting sentence, and four events of varying valence. One Positive Valence scenario had two positive and two neutral events, two Negative Valence scenarios had two negative and two neutral events, and one Conflicting Valence scenario had two positive and two negative events. The varying valence of these events mimicked daily life, where experiences seldom consist of only positive or only negative events, and the four scenarios described four of the six domains targeted in treatment—health, family/friends, evaluations/performance, and finances—without overlapping in content with treatment scenarios. After reading each scenario, participants rated the likelihood of three events' (positive, negative, and neutral) happening next on Likert items ranging from 1 (*very unlikely*) to 7 (*very likely*). These future events were presented in the same random order to all participants and at all assessment points.

The task was given at screening, after Sessions 1-4, and at 1-month follow-up. To assess eligibility at screening, a relative expectancy bias index score was computed by subtracting the mean perceived likelihood of the four negative events (absolute negative bias) from that of the four positive events (absolute positive bias); in this way, the score accounts for expectations of positive and negative events simultaneously. Eligible participants had index scores below a 1.1111 cutoff, determined by subtracting 0.5 standard deviations from the mean (1.65, *SD* = 1.08) relative expectancy bias index scores for 776 college students (see Namaky et al., 2019). This meant that participants' biases were more negative than those of nearly 70% of the prior sample, but not necessarily negative at an absolute level. See Section S1.4 for more details.

To understand the effects of treatment on expectations of positive events and negative events separately, the absolute negative bias and absolute positive bias means were analyzed rather than the relative (difference) index score. Items for the four neutral events were not analyzed. Internal consistency based on McDonald's omega total for the ITT sample using complete item-level data[3] at pretreatment was unacceptable for the negative events, $\omega_t = .28$, 95% CI = [.14, .35], and the positive events, $\omega_t = .31$, 95% CI = [.23, .38], and for the PP sample plausible estimates and stable standard errors did not emerge.[4] We had assumed one dimension for the negative events and another for the positive events, but confirmatory factor analyses for each latent factor with the `OpenMx` package (ver. 2.13.2; Neale et al., 2016) in R showed poor model fit for the ITT and PP samples (see Section S1.6 and Table S4 for details)—a limitation.

Anxiety and depression symptoms. Anxiety and depression symptoms were assessed with the Patient Health Questionnaire-4 (PHQ-4; Kroenke, Spitzer, Williams, & Löwe, 2009), a self-report of core symptoms with four 4-point Likert items ranging from 0 (*not at all*) to 3 (*nearly every day*). The time frame was modified from the past 2 weeks to the past week. Two

---

[3]As stated in the Missing Data Handling section below, across outcome measures only 0.0-0.3% of ITT participants' scale scores were computed from items with at least one item missing. Given this low rate of item-level missingness, the disadvantages of listwise deletion (Enders, 2010, pp. 39-40) did not outweigh its convenience for the purposes of assessing internal consistency.
[4]We computed standardized Cronbach's alpha using the psych package (ver. 1.8.12; Revelle, 2018) in R before learning that methodologists recommend McDonald's omega total instead (Dunn, Baguley, & Brunsden, 2014). We computed omega total using the MBESS package (ver. 4.6.0; Kelley, 2019) in R (see Section S1.5 for details). For transparency, we provide estimates of standardized Cronbach's alpha in Table S3.

anxiety items, identical to those of the Generalized Anxiety Disorder-2 scale (GAD-2; Kroenke, Spitzer, Williams, Monahan, & Löwe, 2007), comprise the Anxiety subscale, for which a sum of 3 or greater reflects potential generalized anxiety, panic, social anxiety, or posttraumatic stress disorder. Two depression items, identical to those of the PHQ-2 (Kroenke, Spitzer, & Williams, 2003), comprise the Depression subscale, for which a sum of 3 or greater reflects potential major depression or another depressive disorder. The PHQ-4 was administered at pretreatment, after Sessions 2 and 4, and at 1-month follow-up. The sum of the anxiety items and the sum of the depression items were analyzed. Internal consistency for the ITT and PP samples using complete item-level data at pretreatment was good for the anxiety items, ω_ts and 95% CIs = .82 [.79, .84] and .83 [.77, .87], and the depression items, ω_ts and 95% CIs = .81 [.78, .84] and .83 [.77, .87].

Self-efficacy, growth mindset, and optimism. Self-efficacy, growth mindset, and optimism were assessed at pretreatment, after Sessions 2 and 4, and at 1-month follow-up using three self-reports. To reduce response burden given high levels of attrition during lengthy assessments in online interventions (e.g., Ji et al., 2019), we identified two or three items per scale using data from Namaky et al. (2019), modified wording, and used a consistent response format—a 5-point Likert scale ranging from 0 (*strongly disagree*) to 4 (*strongly agree*). See Section S1.7 for details. The mean of the selected items for each scale was analyzed.

Self-efficacy was assessed with three items from the New General Self-Efficacy (NGSE) Scale: “When facing difficult tasks, I am certain I will accomplish them” (modified); “I am confident that I can perform effectively on many different tasks”; and “Compared to other people, I can do most tasks very well” (Chen, Gully, & Eden, 2001). Internal consistency for the ITT and PP samples using complete item-level data at pretreatment was good, ω_ts and 95% CIs = .82 [.79, .84] and .83 [.79, .86].

Growth mindset was assessed with three adapted items from a set of growth mindset questions (GMQ) about intelligence (Dweck, 2006): "You can learn new things, but you can't really change how you think" (reverse scored); "No matter how much you have been thinking a particular way, you can always change it quite a bit"; and "You can always substantially change how you think." Internal consistency for the ITT and PP samples using complete item-level data at pretreatment was good, ω_ts and 95% CIs = .80 [.78, .83] and .83 [.78, .86].

Optimism was assessed with two items from the Life Orientation Test-Revised (LOT-R): "If something can go wrong with me, it will" (modified) and "I hardly ever expect things to go my way" (both reverse scored; Scheier, Carver, & Bridges, 1994). Internal consistency for the ITT and PP samples using complete item-level data at pretreatment was good, ω_ts and 95% CIs = .80 [.77, .83] and .81 [.75, .85].

Conditions

At the start of each session participants completed two questions about current positive and negative feelings and then one of the four CBM conditions' tasks or the Neutral Control task (similar to the tasks used in Namaky et al., 2019). Each task consisted of 40 scenarios, and each scenario consisted of three sentences whose outcome resolved when the participant completed a word fragment in the final word or phrase of the third sentence. Each word fragment had one or two missing letters (see below) removed at random when the fragment appeared. After two thirds of the scenarios, participants answered a comprehension question to confirm their understanding of the scenario and reinforce the resolved outcome. See Sections S1.8 and S1.9 for details.

Cognitive bias modification tasks. In the CBM tasks, the 40 scenarios in each session were randomly chosen from a set of 49. Scenarios were not repeated within a session but could be chosen more than once across sessions. The scenarios could end positively or negatively and thus were emotionally ambiguous until they were resolved (e.g., “After being inactive for a few years, you recently joined a recreational soccer league. There is a tournament at the end of the season. You believe that you will contribute to your team’s _____.”). The endings varied by condition. In Positive Prospection, 90% of the scenarios ended positively (e.g., “su_cess”) and 10% ended negatively (e.g., “fai_ure”). In Positive Prospection + Negation, 90% of the scenarios negated a negative outcome in the last sentence and then ended positively (e.g., “. . . You believe that you will *not let your teammates down, and* contribute to your team’s succ_ss.”; emphasis added), and 10% ended negatively. In 50/50 Blocked, 50% of the scenarios ended positively and 50% ended negatively, with the valence alternating every five scenarios. In 50/50 Random, 50% of the scenarios ended positively and 50% ended negatively in random order. To reduce the chances that participants would see the same ending even if the same scenario was used across sessions, for variety, different endings of a given valence were used in Sessions 1 and 3 than in Sessions 2 and 4. The valence of the endings for a given scenario was not fixed across sessions.

To increase desirable difficulty in an effort to increase engagement and learning, the number of missing letters in word fragments varied across sessions. In Sessions 1-2, only one letter was missing (e.g., see above); in Sessions 3-4, two letters were missing, and participants completed each missing letter in turn (e.g., first blank in “vi_to_y,” then second blank in “victo_y”). The type of comprehension question also varied across sessions. In Sessions 1 and 4, participants answered yes/no questions (e.g., “Will your performance probably contribute to the team’s success?”), whereas in Sessions 2-3, participants answered which of two options completed a given sentence (e.g., “Your performance will likely . . .” [a] “help your team win.” or [b] “drag your team down.”). Although the same yes/no question was presented for a scenario if it was chosen at both Sessions 1 and 4, different multiple-choice questions were given for a scenario if it was chosen at Sessions 2 and 3 to increase variety.

Neutral Control task. In the Neutral Control task, the 40 scenarios in each session were randomly chosen from a set of 48. Scenarios were not repeated within a session but could be chosen more than once across sessions. The scenarios, which controlled for the format of CBM but lacked emotional ambiguity about the future, ended neutrally (e.g., "You are in the car with a friend. You think about how it has been a while since you had your car inspected. You decide to get your car inspected next we_k."). As in the CBM tasks, the number of missing letters varied across sessions (one in Sessions 1-2, two in Sessions 3-4). However, the same ending was presented for a given scenario even if the scenario was chosen across sessions, and the type of comprehension question did not vary; participants answered yes/no questions in all sessions, and the same yes/no question was presented for a scenario if it was chosen at multiple sessions.

Implementation. The three sentences in each scenario were presented on the same page one at a time, with each additional sentence after the first appearing once the participant clicked a Continue button. The participant's score, which indicated the number of scenarios for which the participant completed the word fragment correctly on the first attempt, and the participant's number of scenarios completed in the session (out of 40) were displayed at the top of each page. Participants rated how vividly they imagined the scenario after Scenarios 1, 2, and 20, and they rated how vividly they imagined and how much they could relate to all the scenarios on average after Scenario 40. Participants then completed the session's assessment battery and rated if and when they planned to complete the next session. Participants had the ability to view their overall progress through the study on a dashboard page.

Statistical Analysis

All significance tests in this manuscript are two-tailed, and the alpha level is .05 (except in multilevel models, where the Bonferroni-corrected level is .025—see the section Multilevel Modeling). See Section S2 for descriptions of deviations from our preregistered analytic plan.

Baseline differences. We initially tested for baseline demographic differences among the five conditions for ITT participants using unimputed data (see Section S1.10). Only later did we read that randomization ensures that any differences between conditions at baseline are due to chance, making significance tests of baseline differences illogical (Moher et al., 2010). Instead, recommendations are that any covariates should be selected either a priori based on substantial association between covariate and outcome shown in prior research or based on methodological concerns about bias given, for example, nonrandom missing data (Gruijters, 2016). We did not include any demographic covariates a priori (but did include empirically determined auxiliary variables in the imputation model; see Missing Data Handling). Baseline differences in outcomes were not tested because a random intercept was included in each outcome analysis.

Longitudinal outcomes.

Preprocessing. To compare longitudinal outcomes between multiple combinations of conditions, two versions of the dataset with 958 ITT participants were created: a Combined-Level Dataset with the condition variable dummy coded in three levels—Both Positive (including Positive Prospection and Positive Prospection + Negation), Both 50/50 (including 50/50 Blocked and 50/50 Random), and Neutral Control—and a Separate-Level Dataset with the condition variable dummy coded in five levels, one for each of the five conditions.[5] Dummy rather than contrast coding was used because in the Combined-Level Dataset, dummy coding

[5]In the Combined-Level Dataset, condition was coded as 1 (Neutral Control), 2 (Both Positive), and 3 (Both 50/50). In the Separate-Level Dataset, condition was coded as 1 (Neutral Control), 2 (Positive Prospection + Negation), 3 (Positive Prospection), 4 (50/50 Blocked), and 5 (50/50 Random). Although this coding, with Neutral Control as the reference group, was used in the multiple imputation models, in the analysis models the coding was changed based on the effect of interest (e.g., for testing Both Positive against Both 50/50 the coding in the Combined-Level Dataset was changed to 1 [Both 50/50, the reference group], 2 [Both Positive], and 3 [Neutral Control]).

accounts for unequal sample sizes between conditions that are combined into one level (e.g., 177 Positive Prospection and 147 Positive Prospection + Negation participants in Both Positive).

Missing data handling. The present data exhibits two missing data patterns. At the item level, participants had the option to select "prefer not to answer" for most scales (not the PHQ-4), resulting in a general missing data pattern when they did so. This was rare, however; across the seven outcomes, only 0.0-0.3% of ITT participants' scale scores were computed from items with at least one item missing. In such cases, the mean of the available items was analyzed.

At the scale level, attrition yielded a monotone missing data pattern. For ITT participants, the proportions of scale-level missing data across the seven outcomes ranged from 48.5% to 52.2% (see Table S1 for the number of observations of each outcome over time in each condition). To identify measured variables other than time that may relate to this pattern of missing data, we tested whether each demographic variable predicted the number of missing sessions (using nonparametric tests because this number is not normally distributed, but rather negatively skewed). Age and education were the only significant predictors (see Section S1.11 for details). Following an inclusive analysis strategy (Collins, Schafer, & Kam, 2001), we included age and education as auxiliary variables in the multiple imputation model below to correct for any systematic bias resulting from these variables' relationships with missingness.

For the Combined-Level and Separate-Level Datasets, we used the `jomo` (ver. 2.6-8; Quartagno & Carpenter, 2019) and `mitml` packages (ver. 0.3-7; Grund, Robitzsch, & Luedtke, 2019) in R to impute missing scale scores with a joint multivariate linear mixed model. In this multilevel multiple imputation model, the target variables were the seven incomplete Level 1 outcomes (positive and negative expectancy bias, depression, anxiety, self-efficacy, growth mindset, optimism) and two incomplete Level 2 auxiliary variables (age, education). Complete variables by study design were condition and time, for which we included the associated fixed effects of condition, time, and the Condition × Time interaction along with a random intercept and a random slope for time. We treated target variables as continuous and followed Grund, Lüdtke, and Robitzsch (2018) to specify the model and impute 100 datasets. Multivariate normal and missing-at-random data were assumed for the imputation and subsequent analysis models.

Each imputed dataset was then split into two subsets—one subset for the treatment phase, with five assessment points coded as 0 for Baseline and as integers from 1 to 4 for Sessions 1-4, and one subset for the follow-up phase, with two assessment points coded as 0 for Session 4 and as 1 for Follow-Up. By explicitly imputing data before subsetting the datasets by phase rather than implicitly imputing data during analysis (e.g., with maximum likelihood estimation), we ensured that the longitudinal outcome analyses for the different phases were based on the same imputed datasets. Because the PHQ-4, NGSE, GMQ, and LOT-R were not assessed at Sessions 1 or 3, rows containing imputed data at these time points were removed before we analyzed the anxiety, depression, self-efficacy, growth mindset, and optimism outcomes.

Multilevel modeling. We conducted analyses separately for each of the imputed datasets and, following Rubin's rules, pooled the results with the `mitml` package. Because some of the analyses involve small samples (e.g., for the simple effects of time in the PP sample for Positive Prospection + Negation, n = 36), the `df.com` argument of the `testEstimates` function was used to adjust the degrees of freedom with Barnard and Rubin's (1999) procedure and ensure that they would not exceed those had the data been complete. The `confinf` function was used to compute Bonferroni-corrected 97.5% confidence intervals (see below) for final estimates.

Differential change over time between conditions was assessed using hierarchical linear models with restricted maximum likelihood estimation specified using the `nlme` package (ver.

3.1-137; Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2018) in R. We used the `optim` optimizer (Nelder-Mead), which reduced convergence errors we encountered with the default `nlminb` optimizer. We assumed linear trajectories and in each model simultaneously entered fixed effects of condition, time (i.e., assessment point), and the Condition × Time interaction. We conducted separate analyses for treatment and follow-up phases, assuming different trajectories.[6] Treatment phase models included a random intercept and a random slope for time. Follow-up phase models included only a random intercept (they consisted of only two time points).

The Combined-Level Dataset was used to compare (a) the two positive conditions with Neutral Control, (b) the two 50/50 conditions with Neutral Control, and (c) the two positive conditions with the two 50/50 conditions. The Separate-Level Dataset was used to compare (d) Positive Prospection + Negation with Positive Prospection and (e) 50/50 Blocked with 50/50 Random. We coded the latter level of each comparison as the reference group and interpreted the fixed effects of only these five interactions. We did not interpret the other interactions (e.g., Positive Prospection + Negation vs. 50/50 Random in the Separate-Level Dataset) or the lower-order fixed effects. Given that only two of the interactions we interpreted per dataset were orthogonal, an alpha level of .025 (.05 / 2) was used for these analyses and analyses of the simple effects of time. If an interaction was significant, the simple effects of time were assessed at the two condition levels being compared in the interaction, using separate models with a fixed effect of time, a random intercept, and, for treatment phase models, a random slope for time (see Tables 2 and S7 for the ITT and PP samples). Finally, because in the models containing condition and the interaction the main effect of time depends on the condition's reference group, the simple effects of time in all five conditions were assessed in the Separate-Level Dataset to understand

[6]We used separate models for the treatment and follow-up phases assuming different trajectories a priori rather than a piecewise growth curve containing different trajectories in the same model with an empirically derived breakpoint.

the overall change in all conditions, regardless of the interaction's significance (see Tables 1 and S6 for the ITT and PP samples).[7] We did this instead of assessing the main effect of time across all five conditions because main effects are misleading when the interaction is significant.

Effect size. The between-groups effect size for each interpreted interaction was computed as growth-modeling analysis *d* (GMA *d*; Feingold, 2009), which has the same metric as Cohen's *d*, at the end of the treatment phase (i.e., Baseline to Session 4), using the pooled within-group standard deviation at Baseline, and at the end of the follow-up phase (i.e., Session 4 to Follow-Up), using the pooled within-group standard deviation at Session 4. In each case, we multiplied the final parameter estimate for the interaction effect from the pooling phase of the imputation procedure by the number of assessment points following the first assessment point in each study phase (4 for treatment and 1 for follow-up) and divided by the mean of the pooled within-group standard deviations we computed separately in each imputed dataset. The same method was used to compute a 97.5% confidence interval for the GMA *d* based on the upper and lower limits of the 97.5% confidence interval for the final parameter estimate of the interaction effect.

For the within-group effect size of each interpreted simple effect of time, we multiplied the final parameter estimate for the time effect from the imputation pooling phase by the number of assessment points following the first assessment point in each study phase (giving a numerator equal to the difference between the group's model-estimated means at the first and last points of each study phase) and divided by the mean of the group's standard deviation at either Baseline or Session 4 that we computed separately in each imputed dataset. This yields a GMA analogue of effect size for a one-group pretest-posttest design (A. Feingold, personal communication, March

[7]We removed the random slope for time in three of these models for the PP sample in the treatment phase (simple effects of time on self-efficacy in Positive Prospection + Negation and Positive Prospection and on growth mindset in 50/50 Blocked) because they did not converge, leaving only the fixed effect of time and a random intercept.

3, 2019). Confidence intervals were not computed for these effect sizes because equations for the standard errors have not been derived (A. Feingold, personal communication, March 4, 2019).

Iatrogenic effects. We assessed iatrogenic effects on the relative expectancy bias index score (absolute positive bias – absolute negative bias) used to determine eligibility by analyzing the percent decrease on this score from screening to each assessment point based on the raw data (before imputation). Because the index scores could range from –6 to 6, we translated them into the positive range before computing the percent decrease (see Section S1.13 for details). An iatrogenic effect was defined as a decrease of at least 50%. (The program scored this measure in real time and immediately alerted participants that their score had worsened from baseline and offered mental health resources, including service referrals. Links to these resources were also available to all participants via the website's main menu.)

Results

Longitudinal Outcomes

No significantly different changes emerged during treatment or follow-up between the positive conditions (Positive Prospection + Negation vs. Positive Prospection) or between the 50/50 conditions (50/50 Blocked vs. 50/50 Random) for the ITT or PP samples (full results are available in Tables 2 and S7). Therefore, we focus below on differential change for the combined positive and combined 50/50 conditions. Moreover, because the ITT analyses retain groups that systematically differ only by randomized condition, permitting causal inferences about treatment effects, we report only ITT results below (see Altman, 2009, and Hollis & Campbell, 1999). PP analyses revealed similar results, but with fewer significant effects (full results are available in Tables S6 and S7), presumably due to reduced power given the smaller sample and to potentially biased estimates given nonrandom attrition in the full ITT sample.

Positive expectancy bias. In regard to within-group change, ITT participants in all five conditions significantly improved in positive bias during treatment, βs = 0.31 to 0.50, *p*s < .001, *d*s = 1.23 to 2.03, and those in every condition except 50/50 Blocked continued to significantly improve from posttreatment to follow-up, significant βs = 0.25 to 0.51, largest $p = .001$, *d*s = 0.19 to 0.39 (Table 1). Regarding between-group change, as hypothesized, during treatment participants in the two positive conditions improved significantly more than those in Neutral Control, $\beta = 0.19$, $p < .001$, $d = 0.79$, and those in the two 50/50 conditions, $\beta = 0.14$, $p = .001$, $d = 0.58$ (Table 2). However, significantly different changes did not emerge between the two 50/50 conditions and Neutral Control. No significant differences emerged when comparing the change trajectories between conditions from posttreatment to follow-up, indicating that maintenance of treatment gains or further improvement were comparable across conditions.

Negative expectancy bias. Mirroring findings for positive bias, ITT participants in all conditions significantly improved in negative bias during treatment, βs = -0.20 to -0.34, *p*s < .001, *d*s = -0.74 to -1.51, and those in every condition except 50/50 Blocked continued to significantly improve from posttreatment to follow-up, significant βs = -0.18 to -0.33, largest $p = .013$, *d*s = -0.16 to -0.30 (Table 1). As hypothesized, during treatment participants in the two positive conditions improved significantly more than those in Neutral Control, $\beta = -0.14$, $p < .001$, $d = -0.57$, and those in the two 50/50 conditions, $\beta = -0.13$, $p < .001$, $d = -0.55$ (Table 2). Again, significantly different changes did not emerge between the two 50/50 conditions and Neutral Control, nor did any significant differences in trajectories emerge between the conditions we compared from posttreatment to follow-up.

Depression symptoms. ITT participants in all five conditions significantly improved in depression symptoms during treatment, βs = -0.15 to -0.24, largest $p = .009$, *d*s = -0.34 to -0.48, and no condition significantly changed from posttreatment to follow-up (Table 1). Contrary to our hypotheses, no significantly different changes emerged between any of the conditions we compared during treatment or from posttreatment to follow-up (Table 2).

Anxiety symptoms. Paralleling findings for depression symptoms, ITT participants in all five conditions significantly improved in anxiety symptoms during treatment, βs = -0.17 to -0.32, largest p = .002, ds = -0.34 to -0.71 (Table 1), no condition significantly changed from posttreatment to follow-up, and no significantly different changes emerged between any of the conditions we compared during treatment or from posttreatment to follow-up (Table 2).

Self-efficacy. ITT participants in all conditions significantly improved in self-efficacy during treatment, βs = 0.11 to 0.19, ps < .001, ds = 0.50 to 0.85, and only those in Positive Prospection continued to significantly improve from posttreatment to follow-up, β = 0.18, p = .013, d = 0.22 (Table 1). As hypothesized, during treatment participants in positive conditions improved significantly more than those in Neutral Control, β = 0.06, p = .015, d = 0.28, but no significantly different changes emerged between the other conditions we compared during treatment or from posttreatment to follow-up (Table 2).

Growth mindset. Yielding findings similar to those for depression and anxiety, ITT participants in all conditions significantly improved in growth mindset during treatment, βs = 0.07 to 0.14, ps < .001, ds = 0.33 to 0.69 (Table 1), no condition significantly changed from posttreatment to follow-up, and no significantly different changes emerged between any of the conditions we compared during treatment or from posttreatment to follow-up (Table 2).

Optimism. ITT participants in all five conditions significantly improved in optimism during treatment, βs = 0.10 to 0.19, ps < .001, ds = 0.41 to 0.78, and only those in Positive Prospection + Negation continued to significantly improve from posttreatment to follow-up, β =

0.18, p = .022, d = 0.18 (Table 1). As expected, during treatment participants in positive conditions improved significantly more than those in Neutral Control, β = 0.07, p = .009, d = 0.28, but no significantly different changes emerged between the other conditions we compared during treatment or from posttreatment to follow-up (Table 2).

Iatrogenic Effects

One ITT participant's relative expectancy bias index score decreased more than 50% from screening to a later assessment point. This participant, in 50/50 Random, had a decrease of 62.5% from screening (raw score = -2.00, translated score = 6.00) to Session 4 (raw score = -5.75, translated score = 2.25). The participant's positive bias decreased by 2 points (from 3.00 to 1.00) and negative bias increased by 1.75 points (from 5.00 to 6.75). No other participants demonstrated iatrogenic effects based on the established criterion.

Discussion

The present randomized controlled trial evaluated a brief online CBM intervention to train more positive future thinking in a large sample of community participants with relatively negative prospection, a transdiagnostic cognitive process common in emotional disorders. As hypothesized, during treatment ITT participants in the positive conditions improved significantly more in positive and negative expectancy bias, self-efficacy, and optimism than participants in the Neutral Control condition, with the improvements in bias measures also significantly greater than those for participants in the 50/50 conditions. Although ITT participants in all conditions improved on all outcomes during treatment and either did not significantly worsen or continued to significantly improve during follow-up, unexpectedly participants in positive conditions did not improve in anxiety, depression, or growth mindset significantly more than control or 50/50 participants. Moreover, no significantly different changes emerged between Neutral Control and the 50/50 conditions, between the two 50/50 conditions, or between the two positive conditions. PP participants (treatment completers) showed similar results, with fewer significant effects.

Superior Improvement in Expectancy Bias and Positive Outlook

The superior improvements in expectancy bias and two out of the three trait measures of positive outlook in the positive conditions (relative to the control condition) in the present study are broadly consistent with the results of Namaky et al. (2019). In both studies, participants in the positive conditions increased in positive expectancy bias and self-efficacy significantly more than control participants. The present study also found that the positive conditions were superior for increasing optimism, with no evidence of superiority for increasing growth mindset, whereas Namaky et al. found the converse: superiority for increasing growth mindset but not optimism. In addition to finding superiority for increasing positive bias, the present study found superiority for decreasing negative bias, with improvements in the positive conditions surpassing not only those in the control condition but also those in the 50/50 conditions. Together, the studies suggest that resolving ambiguous future scenarios with mostly positive endings improves expectancy bias and self-efficacy, with mixed findings for growth mindset and optimism.

Interestingly, the present study revealed that improvements did not significantly differ between the two positive conditions or between the two 50/50 conditions. The results comparing the positive conditions do not support either the hypothesis that the negation phrase helps by disconfirming negative expectations (Seligman et al., 2013) or the hypothesis that it harms by reinforcing negative associations (Ouimet, Gawronski, & Dozois, 2009). Similarly, the results comparing the 50/50 conditions do not support the hypothesis that shifting future thinking in a given block after first developing a positive or negative expectation in the previous block is more efficacious than shifting future thinking at random points. This may be because the random order gives far more practice shifting between positive and negative outcomes than the blocked order, and this repeated shifting may itself enhance flexibility even when no expectation for a given outcome has been learned, as we assume occurs in the blocked condition, but this is speculative.

The similarity in results between the present study and Namaky et al. (2019) is notable given their methodological differences. For example, whereas Namaky et al. used a small sample of college students receiving course credit at a U.S. university, the present study recruited a large sample of community adults from around the world. Additionally, Namaky et al. used a larger set of scenarios in the Expectancy Bias Task and full-length self-reports, whereas the present study used shorter measures (in most cases by selecting a few representative items) and changed some response scales to facilitate online participation. Several aspects of the CBM conditions also differed (e.g., number of scenarios per session). These differences and others (e.g., assessment points, follow-up period, analyses) may explain some of the minor divergences in results between the studies, but the similar pattern increases our confidence in their shared findings.

Comparable Improvement in Anxiety and Depression Symptoms

The present study and Namaky et al. (2019) are also aligned in finding no evidence of superior improvement in anxiety or depression symptoms; rather, comparable improvement occurred across all conditions. (Although Namaky et al. found that the positive conditions were superior to 50/50 Random for decreasing depression symptoms, we did not find this for the 50/50 conditions in the present study). Of course, because neither study recruited based on symptoms, it may be that symptom improvement should be expected only for participants with sufficient symptom severity at baseline. Given that nearly three fourths of the present sample scored above thresholds for a likely anxiety disorder, depression, or both, analyses on these subgroups may be a starting point for testing this hypothesis (though the analyses will lack randomization to condition within each subgroup and have reduced power).

Another possibility is that the efficacy of the positive training conditions needs further improvement, perhaps through greater personalization of training scenarios. As clinical science moves toward a process-based approach, a focus will be on identifying and targeting specific biopsychosocial processes for individuals with specific goals in specific contexts (Hofmann & Hayes, 2019; Hayes et al., 2019). Although the present training instructs participants to imagine themselves in the scenarios even if the scenarios differ from the participants' usual experiences, more personalized versions might select each person's scenarios from self-reported problem domains, adjust scenario language based on demographic characteristics (e.g., relationship and employment status), or use machine learning to recommend a training set based on participant ratings of a small set of sample scenarios (e.g., Niles & O'Donovan, 2018). Personalization of the scenarios may also improve engagement, potentially reducing treatment dropout.

Still another possible explanation for the lack of differential symptom improvement by condition concerns the brief measures used to assess anxiety and depression symptoms and the analysis of these outcomes at the single-disorder, rather than the transdiagnostic, level (despite the treatment's transdiagnostic design). Interestingly, in an exploratory analysis, when Namaky et al. (2019) examined differential changes in the 14-item sum of the Anxiety and Depression subscales of the Depression Anxiety Stress Scales-21 (Lovibond & Lovibond, 1995) as one dimension (vs. two separate dimensions), they did find that the positive conditions improved significantly more than the neutral condition (N. Namaky, personal communication, September 1, 2019). However, in the present study exploratory analyses of these symptoms combined into one composite measure, the four-item sum of the PHQ-4, revealed comparable improvement across conditions (see Sections S1.14 and S2.9 and Tables S8-S9).

Feasibility of Online Cognitive Bias Modification for Negative Prospection

The present study provides additional evidence that resolving ambiguous scenarios about the future with mostly positive endings in a brief online CBM program requiring no professional contact can shift future thinking and improve outlook in adults with negative expectations about the future. Moreover, this intervention is feasible to implement in the community, and there is little evidence of iatrogenic effects. A large (relative to most treatment studies), transdiagnostic sample of 1,221 adults from 39 countries enrolled in the study for no payment, with 79.5% across conditions starting the first training session, 43.8% starting the second, 32.1% starting the third, 25.0% starting the fourth, and 12.3% starting the 1-month follow-up assessment. Although we seek to reduce treatment dropout and loss to follow-up, the treatment dropout rate is lower than that for many web-based interventions; for example, only 10% of enrolled participants complete a second module at the popular cognitive behavioral website MoodGym (Batterham et al., 2008). For the 958 ITT participants, lower age and education significantly predicted greater attrition. Future research should test whether other measured variables predict attrition (e.g., baseline severity, training credibility, concurrent interventions, self-reported reasons for leaving, device type, usage data); these variables may not only serve as additional auxiliary variables during missing data handling, but also help identify participants at risk of dropping treatment, who may need additional resources to maintain engagement.

Limitations

The present findings must be viewed in light of several limitations relating to treatment dropout, analyses during follow-up, and latent measurement. First, although our ITT analyses retain all randomized participants, because not all of these participants completed their training condition these analyses test the effect of random assignment to condition rather than the effect of completing different conditions, which ultimately may be of interest to clients and providers (Hernán & Robins, 2017; see also Sterne et al., in press, for discussion of ITT and PP analyses).

Second, in the imputation model we assumed linear trajectories from baseline to follow-up (i.e., we used one linear time variable spanning these assessment points), but in the analysis models we assumed piecewise linear trajectories—one in the treatment phase, and one in the follow-up phase (i.e., we conducted separate analyses for each phase). Because more data were observed during treatment than at follow-up, the imputed data at follow-up were likely overly influenced by the trajectory of change during treatment (typically significant improvement), which may have overestimated improvement during the follow-up phase.

Third, the internal consistency of the bias measures was unacceptable for the ITT sample and unstable or implausible for the PP sample, and our assumption of a single dimension for the positive bias items and another for the negative bias items was not supported by confirmatory factor analyses. Because conclusions about a latent construct depend on the associated measure's validity (Flake, Pek, & Hehman, 2017), the bias results should be interpreted with some caution. Although internal consistency for the non-bias measures was good and factor analyses for these would be inappropriate (each has fewer than four items), test-retest reliability and measurement invariance for all measures should be assessed, especially given that we modified items or used only a subset of original items for every scale, meaning that psychometric properties from prior studies may not replicate (Flake et al., 2017). This situation highlights a key challenge of online research: using measures that are brief (given concerns about acceptability and attrition for this delivery model) yet valid to interpret. Although obtaining a similar pattern of results as Namaky et al. (2019) gives us greater confidence in the present findings, ongoing construct validation, a pillar of research on latent constructs, is required.

Conclusion

Targeting transdiagnostic mechanisms of change with technological interventions holds promise for reducing the burden of mental illness on a large scale. The present study is the first to target negative prospection in community adults with a brief online CBM intervention and to show that doing so is feasible at scale, shifts future thinking, and improves positive outlook. Future work is needed to develop more valid measures of expectancy bias. In addition, research that assesses moderation and mediation of treatment effects and that personalizes training may yield insights about mechanisms of change and improve efficacy, thereby ultimately advancing the availability of evidence-based interventions for people with emotional disorders.

References

Altman, D. G. (2009). Missing outcomes in randomized trials: Addressing the dilemma. *Open Medicine, 3*(2), 51-53.

Barnard, J., & Rubin, D. B. (1999). Small-sample degrees of freedom with multiple imputation. *Biometrika, 86*(4), 948-955. doi:10.1093/biomet/86.4.948

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

Batterham, P. J., Neil, A. L., Bennett, K., Griffiths, K. M., & Christensen, H. (2008). Predictors of adherence among community users of a cognitive behavior therapy website. *Patient Preference and Adherence, 2*, 97-105.

Baxter, A. J., Scott, K. M., Vos, T., & Whiteford, H. A. (2013). Global prevalence of anxiety disorders: A systematic review and meta-regression. *Psychological Medicine, 43*(5), 897-910. doi:10.1017/S003329171200147X

Beck, A. T., Brown, G., Steer, R. A., Eidelson, J. I., & Riskind, J. H. (1987). Differentiating anxiety and depression: A test of the cognitive content-specificity hypothesis. *Journal of Abnormal Psychology, 96*(3), 179. doi:10.1037/0021-843X.96.3.179

Blackwell, S. E., Browning, M., Mathews, A., Pictet, A., Welch, J., Davies, J., . . . & Holmes, E. A. (2015). Positive imagery-based cognitive bias modification as a web-based treatment tool for depressed adults: A randomized controlled trial. *Clinical Psychological Science, 3*(1), 91-111. doi:10.1177/2167702614560746

Boland, J., Riggs, K. J., & Anderson, R. J. (2018). A brighter future: The effect of positive episodic simulation on future predictions in non-depressed, moderately dysphoric & highly dysphoric individuals. *Behaviour Research and Therapy, 100*, 7-16.

doi:10.1016/j.brat.2017.10.010

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods, 4*(1), 62-83. doi:10.1177%2F109442810141004

Chisholm, D., Sweeny, 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. doi:10.1016/S2215-0366(16)30024-4

Collins, L. M., Schafer, J. L., & Kam, C-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6*, 330-351. doi:10.1037/1082-989X.6.4.330

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology, 105*, 399-412. doi:10.1111/bjop.12046

Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York, NY: Random House Inc.

Enders, C. K. (2010). *Applied missing data analysis.* New York, NY: Guilford press.

Eysenbach, G. (2005). The law of attrition. *Journal of Medical Internet Research, 7*(1), 1-9. doi:10.2196/jmir.7.1.e11

Feingold, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods, 14*(1), 43-53. doi:10.1037/a0014699

Ferrari, A. J., Somerville, A. J., Baxter, A. J., Norman, R., Patten, S. B., Vos, T., & Whiteford, H. A. (2013). Global variation in the prevalence and incidence of major depressive disorder: A systematic review of the epidemiological literature. *Psychological Medicine,* *43*(3), 471-481. doi:10.1017/S0033291712001511

Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality Science, 8*(4), 370-378.

Gilbert, D. T., & Wilson, T. D. (2007). Prospection: Experiencing the future. *Science, 317*(5843), 1351-1354. doi:10.1126/science.1144161

Gruijters, S. L. K. (2016). Baseline comparisons and covariate fishing: Bad statistical habits we should have broken yesterday. *The European Health Psychologist, 18*(5), 205-209.

Grund, S., Lüdtke, O., & Robitzsch, A. (2016). Multiple imputation of multilevel missing data: An introduction to the R package pan. *SAGE Open*, 1-17. doi:10.1177%2F2158244016668220

Grund, S., Lüdtke, O., & Robitzsch, A. (2018). Multiple imputation of missing data for multilevel models: Simulations and recommendations. *Organizational Research Methods, 21*(1), 111-149. doi:10.1177%2F1094428117703686

Grund, S., Robitzsch, A., & Luedtke, O. (2019). mitml: Tools for Multiple Imputation in Multilevel Modeling [Computer software]. Retrieved from https://cran.r-project.org/package=mitml

Hayes, S. C., Hofmann, S. G., Stanton, C. E., Carpenter, J. K., Sanford, B. T., Curtiss, J. E., & Ciarrochi, J. (2019). The role of the individual in the coming era of process-based therapy. *Behaviour Research and Therapy, 117*, 40-53. doi:10.1016/j.brat.2018.10.005

Hernán, M. A., & Robins, J. M. (2017). Per-protocol analyses of pragmatic trials. *New England Journal of Medicine, 377*(14), 1391-1398. doi:10.1056/NEJMsm1605385

Hofmann, S. G., & Hayes, S. C. (2019). The future of intervention science: Process-based therapy. *Clinical Psychological Science, 7*(1), 37-50. doi:10.1177%2F2167702618772296

Hollis, S., & Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomized controlled trials. *BMJ, 319*, 670-674. doi:10.1136/bmj.319.7211.670

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. doi:10.1080/10705519909540118

Ji, J. L., Baee, S., Zhang, D., Meyer, M. J., Funk, D., Portnow, S., Barnes, L., & Teachman, B. A. (2019). *Multi-session online interpretation bias training for anxiety in a community sample.* Manuscript submitted for publication.

Jones, E. B., & Sharpe, L. (2017). Cognitive bias modification: A review of meta-analyses. *Journal of Affective Disorders, 223*, 175-183. doi:10.1016/j.jad.2017.07.034

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1-27. doi:10.1146/annurev.clinpsy.3.022806.091432

Kazdin, A. E. (2017). Addressing the treatment gap: A key challenge for extending evidence-based psychosocial interventions. *Behaviour Research and Therapy, 88*, 7-18. doi:10.1016/j.brat.2016.06.004

Kelley, K. (2019). MBESS: The MBESS R Package [Computer software]. Retrieved from https://cran.r-project.org/package=MBESS

Kroenke, K., Baye, F., & Lourens, S. G. (2019). Comparative validity and responsiveness of PHQ-ADS and other composite anxiety-depression measures. *Journal of Affective Disorders, 246*, 437-443. doi:10.1016/j.jad.2018.12.098

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2003). The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Medical Care, 41*(11), 1284-1292. doi:10.1097/01.MLR.0000093487.78664.3C

Kroenke, K., Spitzer, R. L., Williams, J. B., & Löwe, B. (2009). An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics, 50*(6), 613-621. doi:10.1016/S0033-3182(09)70864-3

Kroenke, K., Spitzer, R. L., Williams, J. B. W., Monahan, P. O., & Löwe, B. (2007). Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine, 146*, 317-325. doi:10.7326/0003-4819-146-5-200703060-00004

Kroenke, K., Wu, J., Yu, Z., Bair, M., Kean, J., Stump, T., & Monahan, P. (2016). Patient Health Questionnaire Anxiety and Depression Scale: Initial validation in three clinical trials. *Psychosomatic Medicine, 78*(6), 716-727. doi:10.1097/PSY.0000000000000322

Lang, T. J., Blackwell, S. E., Harmer, C. J., Davison, P., & Holmes, E. A. (2012). Cognitive bias modification using mental imagery for depression: Developing a novel computerized intervention to change negative thinking styles. *European Journal of Personality, 26*(2), 145-157. doi:10.1002/per.855

Lovibond, S.H. & Lovibond, P.F. (1995). *Manual for the Depression Anxiety Stress Scales* (2nd ed.). Sydney: Psychology Foundation.

Löwe, B., Wahl, I., Rose, M., Spitzer, C., Glaesmer, H., Wingenfeld, K., . . . Brähler, E. (2010). A 4-item measure of depression and anxiety: Validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *Journal of Affective Disorders, 122*, 86-95. doi:10.1016/j.jad.2009.06.019

Malouff, J. M., & Schutte, N. S. (2017). Can psychological interventions increase optimism? A meta-analysis. *The Journal of Positive Psychology, 12*(6), 594-604. doi:10.1080/17439760.2016.1221122

Mathews, A., & Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *Journal of Abnormal Psychology, 109*(4), 602-615. doi:10.1037/0021-843X.109.4.602

Meevissen, Y. M., Peters, M. L., & Alberts, H. J. (2011). Become more optimistic by imagining a best possible self: Effects of a two week intervention. *Journal of Behavior Therapy and Experimental Psychiatry, 42*(3), 371-378. doi:10.1016/j.jbtep.2011.02.012

Menne-Lothmann, C., Viechtbauer, W., Höhn, P., Kasanova, Z., Haller, S. P., Drukker, M., . . . & Lau, J. Y. (2014). How to boost positive interpretations? A meta-analysis of the effectiveness of cognitive bias modification for interpretation. *PLoS One*(6), e100925. doi:10.1371/journal.pone.0100925

Miloyan, B., Pachana, N. A., & Suddendorf, T. (2014). The future is here: A review of foresight systems in anxiety and depression. *Cognition and Emotion, 28*(5), 795-810. doi:10.1080/02699931.2013.863179

Miloyan, B., & Suddendorf, T. (2015). Feelings of the future. *Trends in Cognitive Sciences, 19*(4), 196-200. doi:10.1016/j.tics.2015.01.008

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., . . . Altman, D. G. (2010). CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. *BMJ, 340*, 1-28. doi:10.1136/bmj.c869

Murphy, S. E., O'Donoghue, M. C., Drazich, E. H., Blackwell, S. E., Nobre, A. C., & Holmes, E. A. (2015). Imagining a brighter future: The effect of positive imagery training on mood, prospective mental imagery and emotional bias in older adults. *Psychiatry Research, 230*(1), 36-43. doi:10.1016/j.psychres.2015.07.059

Namaky, N., Glenn, J. J., Eberle, J. W., & Teachman, B. A. (2019). *Adapting cognitive bias modification to train healthy prospection.* Manuscript submitted for publication.

Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., . . . Boker, S. M. (2016). OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika, 81*(2), 535-549.

Niles, A. N., & O'Donovan, A. (2018). Personalizing affective stimuli using a recommender algorithm: An example with threatening words for trauma exposed populations. *Cognitive Therapy and Research, 42*, 747-757. doi:10.1007/s10608-018-9923-8

Ouimet, A. J., Gawronski, B., & Dozois, D. J. (2009). Cognitive vulnerability to anxiety: A review and an integrative model. *Clinical Psychology Review, 29*(6), 459. doi:10.1016/j.cpr.2009.05.004

Pham, L. B., & Taylor, S. E. (1999). From thought to action: Effects of process- versus outcome-based mental simulations on performance. *Personality and Social Psychology Bulletin, 25*(2), 250-260. doi:10.1177%2F0146167299025002010

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team (2018). nlme: Linear and Nonlinear Mixed Effects Models. R package [Computer software]. Retrieved from https://CRAN.R-project.org/package=nlme

Quartagno, M., & Carpenter, J. (2019). jomo: A package for multilevel joint modelling multiple imputation [Computer software]. Retrieved from https://cran.r-project.org/package=jomo

Quoidbach, J., Wood, A. M., & Hansenne, M. (2009). Back to the future: The effect of daily practice of mental time travel into the future on happiness and anxiety. *The Journal of Positive Psychology, 4*(5), 349-355. doi:10.1080/17439760902992365

R Core Team (2018). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment, 92*(6), 544-559. doi:10.1080/00223891.2010.496477

Renner, F., Ji, J. L., Pictet, A., Holmes, E. A., & Blackwell, S. E. (2017). Effects of engaging in repeated mental imagery of future positive events on behavioural activation in individuals with major depressive disorder. *Cognitive Therapy and Research, 41*(3), 369-380. doi:10.1007/s10608-016-9776-y

Revelle, W. (2018). psych: Procedures for Psychological, Psychometric, and Personality Research [Computer software]. Retrieved from https://cran.r-project.org/package=psych

Roepke, A. M., & Seligman, M. E. P. (2016). Depression and prospection. *British Journal of Clinical Psychology, 55*(1), 23-48. doi:10.1111/bjc.12087

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A re-evaluation of the Life Orientation Test. *Journal of Personality and Social Psychology, 67*, 1063-1078. doi:10.1037/0022-3514.67.6.1063

Seligman, M. E., Railton, P., Baumeister, R. F., & Sripada, C. (2013). Navigating into the future or driven by the past. *Perspectives on Psychological Science, 8*(2), 119-141. doi:10.1177%2F1745691612474317

Signorell, A. (2019). DescTools: Tools for descriptive statistics. R package [Computer software]. Retrieved from https://cran.r-project.org/package=DescTools

Szpunar, K. K., & Schacter, D. L. (2013). Get real: Effects of repeated simulation and emotion on the perceived plausibility of future experiences. *Journal of Experimental Psychology: General, 142*(2), 323-327. doi:10.1037/a0028877

Szpunar, K. K., Spreng, R. N., & Schacter, D. L. (2014). A taxonomy of prospection: Introducing an organizational framework for future-oriented cognition. *Proceedings of the National Academy of Sciences, 111*(52), 18414-18421. doi:10.1073/pnas.1417144111

Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., . . . Higgins, J. P. T. (in press). RoB 2: A revised tool for assessing risk of bias in randomized trials. *BMJ*. https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool?authuser=0

Taylor, S. E., Pham, L. B., Rivkin, I. D., & Armor, D. A. (1998). Harnessing the imagination: Mental simulation, self-regulation, and coping. *American Psychologist, 53*(4), 429-439. doi:10.1037/0003-066X.53.4.429

Teachman, B. A. (2014). No appointment necessary: Treating mental illness outside the therapist’s office. *Perspectives on Psychological Science, 9*(1), 85-87. doi:10.1177%2F1745691613512659

Tofighi, D., & MacKinnon, D. P. (2011). RMediation: An R package for mediation analysis confidence intervals. *Behavior Research Methods, 43*(3), 692-700. doi:10.3758%2Fs13428-011-0076-x

Wang, P. S., Angermeyer, M., Borges, G., Bruffaerts, R., Chiu, W. T., De Girolamo, G., . . . Kessler, R. C. (2007). Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry, 6*(3), 177.

Table 1

Multilevel Modeling Time Effects in Each Condition for the Intent-To-Treat Sample

Outcome	Phase	Condition	β (*SE*)	*df*	*t*	*p*	*d*
Positive Bias	TX	Positive + Negation	0.50 (0.04)	123.09	11.49	< .001[§§]	2.03
		Positive	0.50 (0.03)	157.78	14.66	< .001[§§]	2.09
		50/50 Blocked	0.32 (0.04)	112.31	7.49	< .001[§§]	1.45
		50/50 Random	0.41 (0.04)	153.65	10.81	< .001[§§]	1.66
		Neutral Control	0.31 (0.03)	224.95	12.28	< .001[§§]	1.23
	FU	Positive + Negation	0.51 (0.13)	53.24	3.98	< .001[§§]	0.39
		Positive	0.45 (0.11)	72.17	4.07	< .001[§§]	0.38
		50/50 Blocked	0.23 (0.13)	52.04	1.78	.081	0.19
		50/50 Random	0.42 (0.11)	79.04	3.85	< .001[§§]	0.33
		Neutral Control	0.25 (0.08)	131.49	3.26	.001[§§]	0.20
Negative Bias	TX	Positive + Negation	-0.33 (0.04)	113.18	-8.67	< .001[§§]	-1.30
		Positive	-0.34 (0.03)	131.15	-10.62	< .001[§§]	-1.51
		50/50 Blocked	-0.20 (0.04)	137.35	-5.65	< .001[§§]	-0.74
		50/50 Random	-0.21 (0.03)	165.01	-6.96	< .001[§§]	-0.94
		Neutral Control	-0.20 (0.02)	213.55	-9.13	< .001[§§]	-0.86
	FU	Positive + Negation	-0.32 (0.11)	58.38	-2.86	.006[§]	-0.27
		Positive	-0.33 (0.10)	73.45	-3.34	.001[§§]	-0.30
		50/50 Blocked	-0.16 (0.11)	60.72	-1.41	.163	-0.13
		50/50 Random	-0.28 (0.10)	75.52	-2.73	.008[§]	-0.23
		Neutral Control	-0.18 (0.07)	126.10	-2.51	.013[§]	-0.16
Anxiety	TX	Positive + Negation	-0.32 (0.06)	80.39	-5.26	< .001[§§]	-0.71
		Positive	-0.17 (0.05)	101.75	-3.22	.002[§§]	-0.34
		50/50 Blocked	-0.18 (0.05)	104.34	-3.44	.001[§§]	-0.42
		50/50 Random	-0.17 (0.05)	110.74	-3.16	.002[§§]	-0.34
		Neutral Control	-0.21 (0.03)	209.41	-6.10	< .001[§§]	-0.44
	FU	Positive + Negation	-0.28 (0.19)	59.31	-1.48	.143	-0.16
		Positive	-0.15 (0.18)	74.76	-0.83	.408	-0.08
		50/50 Blocked	-0.09 (0.20)	67.81	-0.43	.668	-0.05
		50/50 Random	-0.29 (0.18)	76.87	-1.61	.112	-0.16
		Neutral Control	-0.15 (0.13)	127.15	-1.20	.231	-0.09
Depression	TX	Positive + Negation	-0.23 (0.06)	88.10	-3.93	< .001[§§]	-0.47
		Positive	-0.18 (0.05)	89.44	-3.51	.001[§§]	-0.37
		50/50 Blocked	-0.15 (0.06)	88.31	-2.67	.009[§]	-0.34
		50/50 Random	-0.24 (0.05)	102.48	-4.47	< .001[§§]	-0.48
		Neutral Control	-0.20 (0.04)	173.30	-5.86	< .001[§§]	-0.42
	FU	Positive + Negation	-0.18 (0.18)	62.74	-0.99	.326	-0.10
		Positive	-0.15 (0.17)	74.39	-0.91	.367	-0.08
		50/50 Blocked	-0.13 (0.18)	69.14	-0.70	.487	-0.07
		50/50 Random	-0.23 (0.18)	76.03	-1.29	.201	-0.12
		Neutral Control	-0.14 (0.12)	129.79	-1.16	.250	-0.07
Self-Efficacy	TX	Positive + Negation	0.16 (0.03)	76.13	6.02	< .001[§§]	0.76
		Positive	0.19 (0.03)	90.41	7.62	< .001[§§]	0.85

		50/50 Blocked	0.13 (0.03)	84.47	4.68	< .001§§	0.57
		50/50 Random	0.15 (0.02)	131.12	6.20	< .001§§	0.63
		Neutral Control	0.11 (0.02)	142.83	6.53	< .001§§	0.50
	FU	Positive + Negation	0.15 (0.08)	51.37	1.89	.065	0.17
		Positive	0.18 (0.07)	70.03	2.54	.013§	0.22
		50/50 Blocked	0.12 (0.08)	65.58	1.60	.115	0.13
		50/50 Random	0.10 (0.07)	76.82	1.40	.166	0.12
		Neutral Control	0.11 (0.05)	109.81	2.12	.036	0.12
Growth Mindset	TX	Positive + Negation	0.14 (0.03)	77.29	4.50	< .001§§	0.69
		Positive	0.11 (0.03)	83.48	4.26	< .001§§	0.50
		50/50 Blocked	0.11 (0.03)	73.49	3.79	< .001§§	0.54
		50/50 Random	0.14 (0.03)	90.57	5.05	< .001§§	0.57
		Neutral Control	0.07 (0.02)	132.47	3.80	< .001§§	0.33
	FU	Positive + Negation	0.17 (0.08)	56.81	2.05	.045	0.18
		Positive	0.12 (0.08)	71.22	1.62	.109	0.12
		50/50 Blocked	0.14 (0.08)	61.56	1.69	.097	0.15
		50/50 Random	0.10 (0.08)	68.40	1.26	.213	0.10
		Neutral Control	0.07 (0.05)	109.03	1.25	.214	0.07
Optimism	TX	Positive + Negation	0.19 (0.03)	58.41	5.71	< .001§§	0.78
		Positive	0.14 (0.03)	85.62	5.49	< .001§§	0.64
		50/50 Blocked	0.13 (0.03)	87.37	4.85	< .001§§	0.53
		50/50 Random	0.10 (0.03)	107.23	3.79	< .001§§	0.41
		Neutral Control	0.10 (0.02)	182.59	6.06	< .001§§	0.41
	FU	Positive + Negation	0.18 (0.08)	58.65	2.36	.022§	0.18
		Positive	0.17 (0.08)	57.03	2.19	.033	0.17
		50/50 Blocked	0.15 (0.08)	62.22	1.82	.074	0.15
		50/50 Random	0.15 (0.07)	85.50	2.16	.034	0.14
		Neutral Control	0.07 (0.05)	114.14	1.31	.192	0.07

Note. Separate models were fit for each outcome, phase, and condition. Every model included the fixed effect of time (shown here). Treatment phase models included a random intercept and random slope for time; follow-up phase models included only a random intercept. The Separate-Level Dataset, with condition coded in five levels (Positive Prospection + Negation, Positive Prospection, 50/50 Blocked, 50/50 Random, Neutral Control), was used. To correct for multiple comparisons among models in Table 2, the Bonferroni-corrected alpha level is .025. TX = treatment; FU = follow-up.
§$p < .025$
§§$p < .005$

Table 2

Multilevel Modeling Fixed Condition × Time Interaction and Simple Time Effects for the Intent-To-Treat Sample

Outcome	Phase	Effect	β (*SE*)	*df*	*t*	*p*	*d*, 97.5% CI
Positive Bias	TX	(Both Positive vs. Neutral Control) × Time	0.19 (0.04)	223.21	4.87	< .001§§	0.79 [0.42, 1.15]
		$\text{Time}_{\text{Both Positive}}$	0.50 (0.03)	147.94	16.72	< .001§§	2.09
		$\text{Time}_{\text{Neutral Control}}$	0.31 (0.03)	224.95	12.28	< .001§§	1.23
		(Both Positive vs. Both 50/50) × Time	0.14 (0.04)	233.23	3.54	< .001§§	0.58 [0.21, 0.95]
		$\text{Time}_{\text{Both Positive}}$	0.50 (0.03)	147.94	16.72	< .001§§	2.09
		$\text{Time}_{\text{Both 50/50}}$	0.36 (0.03)	203.27	13.68	< .001§§	1.56
		(Both 50/50 vs. Neutral Control) × Time	0.06 (0.04)	344.36	1.62	.106	0.23 [-0.09, 0.55]
		(Positive + Negation vs. Positive) × Time	0.00 (0.05)	272.49	0.08	.940	0.02 [-0.47, 0.50]
		(50/50 Blocked vs. 50/50 Random) × Time	-0.09 (0.06)	225.35	-1.63	.104	-0.39 [-0.94, 0.15]
	FU	(Both Positive vs. Neutral Control) × Time	0.22 (0.11)	289.47	2.08	.039	0.18 [-0.01, 0.37]
		(Both Positive vs. Both 50/50) × Time	0.13 (0.11)	257.93	1.22	.222	0.11 [-0.09, 0.30]
		(Both 50/50 vs. Neutral Control) × Time	0.09 (0.11)	248.74	0.78	.435	0.07 [-0.13, 0.27]
		(Positive + Negation vs. Positive) × Time	0.06 (0.17)	187.23	0.35	.727	0.05 [-0.26, 0.36]
		(50/50 Blocked vs. 50/50 Random) × Time	-0.20 (0.17)	203.10	-1.16	.248	-0.15 [-0.45, 0.15]
Negative Bias	TX	(Both Positive vs. Neutral Control) × Time	-0.14 (0.03)	282.30	-4.25	< .001§§	-0.57 [-0.87, -0.27]
		$\text{Time}_{\text{Both Positive}}$	-0.34 (0.02)	176.03	-14.08	< .001§§	-1.42
		$\text{Time}_{\text{Neutral Control}}$	-0.20 (0.02)	213.55	-9.13	< .001§§	-0.86
		(Both Positive vs. Both 50/50) × Time	-0.13 (0.03)	251.44	-4.03	< .001§§	-0.55 [-0.85, -0.24]
		$\text{Time}_{\text{Both Positive}}$	-0.34 (0.02)	176.03	-14.08	< .001§§	-1.42
		$\text{Time}_{\text{Both 50/50}}$	-0.21 (0.02)	190.47	-8.66	< .001§§	-0.84
		(Both 50/50 vs. Neutral Control) × Time	0.00 (0.03)	286.53	-0.07	.943	-0.01 [-0.31, 0.29]
		(Positive + Negation vs. Positive) × Time	0.01 (0.05)	284.28	0.22	.823	0.04 [-0.38, 0.46]
		(50/50 Blocked vs. 50/50 Random) × Time	0.01 (0.05)	255.90	0.30	.764	0.06 [-0.37, 0.48]
	FU	(Both Positive vs. Neutral Control) × Time	-0.14 (0.11)	212.58	-1.31	.190	-0.12 [-0.33, 0.09]
		(Both Positive vs. Both 50/50) × Time	-0.10 (0.11)	201.31	-0.88	.381	-0.08 [-0.29, 0.13]
		(Both 50/50 vs. Neutral Control) × Time	-0.05 (0.10)	256.33	-0.45	.657	-0.04 [-0.23, 0.16]
		(Positive + Negation vs. Positive) × Time	0.01 (0.16)	191.45	0.06	.950	0.01 [-0.30, 0.32]
		(50/50 Blocked vs. 50/50 Random) × Time	0.12 (0.15)	210.23	0.80	.424	0.10 [-0.18, 0.39]
Anxiety	TX	(Both Positive vs. Neutral Control) × Time	-0.03 (0.05)	252.51	-0.52	.606	-0.06 [-0.31, 0.20]

		(Both Positive vs. Both 50/50) × Time	-0.07 (0.06)	203.09	-1.15	.252	-0.14 [-0.42, 0.14]
		(Both 50/50 vs. Neutral Control) × Time	0.04 (0.05)	277.32	0.74	.459	0.08 [-0.17, 0.33]
		(Positive + Negation vs. Positive) × Time	-0.15 (0.08)	213.18	-1.87	.062	-0.32 [-0.70, 0.07]
		(50/50 Blocked vs. 50/50 Random) × Time	-0.01 (0.08)	235.46	-0.17	.863	-0.03 [-0.40, 0.34]
	FU	(Both Positive vs. Neutral Control) × Time	-0.05 (0.19)	210.75	-0.25	.805	-0.03 [-0.27, 0.22]
		(Both Positive vs. Both 50/50) × Time	0.00 (0.18)	238.63	-0.01	.996	0.00 [-0.23, 0.23]
		(Both 50/50 vs. Neutral Control) × Time	-0.05 (0.18)	262.10	-0.26	.797	-0.03 [-0.25, 0.20]
		(Positive + Negation vs. Positive) × Time	-0.13 (0.26)	237.06	-0.52	.607	-0.07 [-0.40, 0.25]
		(50/50 Blocked vs. 50/50 Random) × Time	0.21 (0.27)	227.32	0.78	.435	0.12 [-0.22, 0.45]
Depression	TX	(Both Positive vs. Neutral Control) × Time	0.01 (0.06)	204.83	0.15	.885	0.02 [-0.24, 0.27]
		(Both Positive vs. Both 50/50) × Time	0.00 (0.05)	230.51	0.05	.958	0.01 [-0.24, 0.25]
		(Both 50/50 vs. Neutral Control) × Time	0.01 (0.05)	241.65	0.10	.920	0.01 [-0.23, 0.26]
		(Positive + Negation vs. Positive) × Time	-0.05 (0.08)	221.33	-0.63	.531	-0.10 [-0.45, 0.26]
		(50/50 Blocked vs. 50/50 Random) × Time	0.09 (0.08)	211.30	1.19	.234	0.20 [-0.18, 0.57]
	FU	(Both Positive vs. Neutral Control) × Time	-0.02 (0.18)	227.74	-0.08	.933	-0.01 [-0.23, 0.21]
		(Both Positive vs. Both 50/50) × Time	0.02 (0.18)	205.44	0.09	.932	0.01 [-0.22, 0.23]
		(Both 50/50 vs. Neutral Control) × Time	-0.03 (0.18)	245.66	-0.18	.861	-0.02 [-0.23, 0.20]
		(Positive + Negation vs. Positive) × Time	-0.02 (0.25)	225.05	-0.09	.928	-0.01 [-0.32, 0.29]
		(50/50 Blocked vs. 50/50 Random) × Time	0.10 (0.25)	240.89	0.40	.687	0.05 [-0.25, 0.35]
Self-Efficacy	TX	(Both Positive vs. Neutral Control) × Time	0.06 (0.03)	224.60	2.46	.015§	0.28 [0.02, 0.53]
		$\text{Time}_{\text{Both Positive}}$	0.18 (0.02)	117.05	9.34	< .001§§	0.81
		$\text{Time}_{\text{Neutral Control}}$	0.11 (0.02)	142.83	6.53	< .001§§	0.50
		(Both Positive vs. Both 50/50) × Time	0.04 (0.02)	261.29	1.47	.142	0.16 [-0.08, 0.39]
		(Both 50/50 vs. Neutral Control) × Time	0.03 (0.03)	233.58	1.08	.283	0.12 [-0.13, 0.36]
		(Positive + Negation vs. Positive) × Time	-0.03 (0.04)	192.97	-0.68	.499	-0.12 [-0.51, 0.28]
		(50/50 Blocked vs. 50/50 Random) × Time	-0.02 (0.04)	188.69	-0.50	.615	-0.08 [-0.46, 0.29]
	FU	(Both Positive vs. Neutral Control) × Time	0.05 (0.08)	192.79	0.60	.553	0.05 [-0.15, 0.25]
		(Both Positive vs. Both 50/50) × Time	0.04 (0.08)	195.92	0.55	.581	0.05 [-0.15, 0.24]
		(Both 50/50 vs. Neutral Control) × Time	0.00 (0.07)	233.07	0.05	.960	0.00 [-0.18, 0.19]
		(Positive + Negation vs. Positive) × Time	-0.03 (0.11)	209.56	-0.25	.801	-0.03 [-0.31, 0.24]
		(50/50 Blocked vs. 50/50 Random) × Time	0.02 (0.10)	266.17	0.17	.867	0.02 [-0.23, 0.27]
Growth Mindset	TX	(Both Positive vs. Neutral Control) × Time	0.06 (0.03)	195.55	2.02	.045	0.26 [-0.03, 0.54]
		(Both Positive vs. Both 50/50) × Time	0.00 (0.03)	206.67	0.10	.922	0.01 [-0.26, 0.29]

		(Both 50/50 vs. Neutral Control) × Time	0.05 (0.03)	239.19	2.07	.040	0.24 [-0.02, 0.50]
		(Positive + Negation vs. Positive) × Time	0.03 (0.04)	179.11	0.75	.452	0.14 [-0.28, 0.56]
		(50/50 Blocked vs. 50/50 Random) × Time	-0.03 (0.04)	183.95	-0.63	.530	-0.11 [-0.52, 0.30]
	FU	(Both Positive vs. Neutral Control) × Time	0.08 (0.08)	228.12	1.06	.290	0.08 [-0.09, 0.26]
		(Both Positive vs. Both 50/50) × Time	0.04 (0.08)	222.18	0.46	.645	0.04 [-0.14, 0.21]
		(Both 50/50 vs. Neutral Control) × Time	0.05 (0.08)	211.42	0.58	.562	0.05 [-0.14, 0.23]
		(Positive + Negation vs. Positive) × Time	0.05 (0.11)	197.76	0.44	.663	0.05 [-0.21, 0.31]
		(50/50 Blocked vs. 50/50 Random) × Time	0.04 (0.11)	251.38	0.34	.737	0.04 [-0.21, 0.29]
Optimism	TX	(Both Positive vs. Neutral Control) × Time	0.07 (0.03)	243.73	2.62	.009[§]	0.28 [0.04, 0.51]
		$\text{Time}_{\text{Both Positive}}$	0.17 (0.02)	112.71	8.26	< .001[§§]	0.72
		$\text{Time}_{\text{Neutral Control}}$	0.10 (0.02)	182.59	6.06	< .001[§§]	0.41
		(Both Positive vs. Both 50/50) × Time	0.05 (0.03)	182.76	1.95	.053	0.23 [-0.04, 0.49]
		(Both 50/50 vs. Neutral Control) × Time	0.01 (0.03)	256.02	0.45	.651	0.05 [-0.18, 0.27]
		(Positive + Negation vs. Positive) × Time	0.04 (0.04)	179.53	1.11	.269	0.19 [-0.20, 0.58]
		(50/50 Blocked vs. 50/50 Random) × Time	0.03 (0.04)	202.99	0.76	.447	0.12 [-0.23, 0.47]
	FU	(Both Positive vs. Neutral Control) × Time	0.10 (0.08)	209.38	1.27	.207	0.10 [-0.08, 0.27]
		(Both Positive vs. Both 50/50) × Time	0.01 (0.08)	203.67	0.19	.853	0.01 [-0.16, 0.19]
		(Both 50/50 vs. Neutral Control) × Time	0.08 (0.08)	207.76	1.07	.285	0.08 [-0.09, 0.25]
		(Positive + Negation vs. Positive) × Time	0.01 (0.11)	176.93	0.12	.906	0.01 [-0.24, 0.27]
		(50/50 Blocked vs. 50/50 Random) × Time	0.00 (0.10)	240.12	-0.03	.979	0.00 [-0.23, 0.22]

Note. Separate models were fit for each outcome, phase, and reference group. Each model contained the fixed effects of condition, time, and the Condition × Time interaction. Treatment phase models included a random intercept and random slope for time; follow-up phase models included only a random intercept. The latter level of the dummy-coded condition factor in each interaction effect is the reference group. Simple time effects are shown only for significant interactions. Significance is based on a Bonferroni-corrected alpha level of .025 (.05/2 given two orthogonal interactions per dataset). The Combined-Level Dataset, with condition coded in three levels (Both Positive, Both 50/50, Neutral Control), was used to test interactions contrasting Both Positive with Neutral Control, Both Positive with Both 50/50, and Both 50/50 with Neutral Control. The Separate-Level Dataset, with condition coded in five levels (Positive Prospection + Negation, Positive Prospection, 50/50 Blocked, 50/50 Random, Neutral Control), was used to test interaction effects contrasting Positive Prospection + Negation with Positive Prospection and 50/50 Blocked with 50/50 Random. TX = treatment; FU = follow-up.

[§]$p < .025$

[§§]$p < .005$

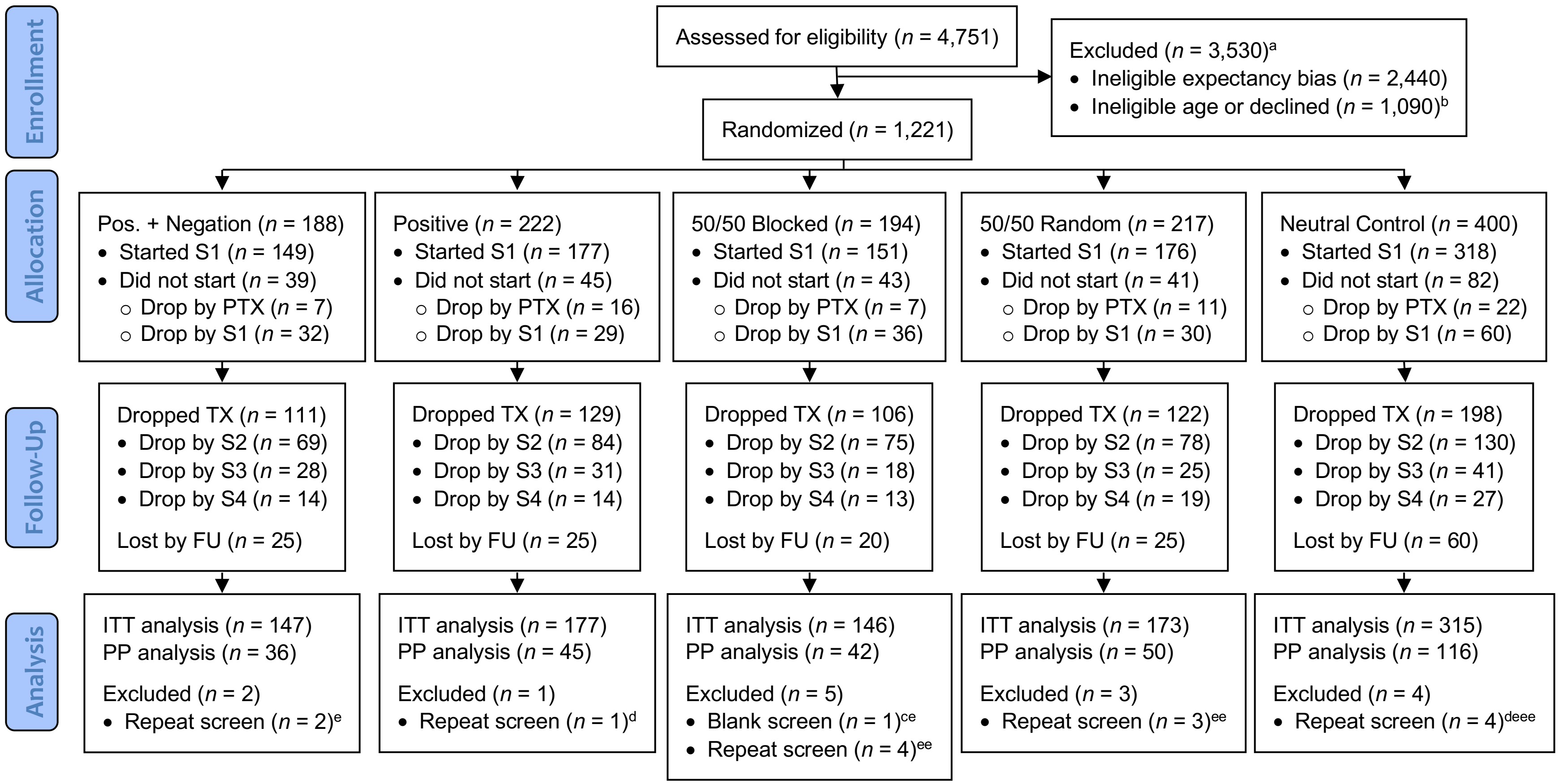

Figure 1. Participant flowchart. Numbers *dropped by* a given session reflect participants who did not start (vs. started but did not complete) the session. Numbers *lost by* follow-up reflect participants who did not start (vs. started but did not complete) follow-up. S1-S4 = Session 1-Session 4; PTX = pretreatment; TX = treatment; FU = follow-up; ITT = intent-to-treat (started S1); PP = per-protocol (completed S4).

[a]May include multiple screening attempts by the same participants for participants whose browser cookies were disabled.
[b]Website design did not differentiate participants under 18 years of age from eligible participants who declined to enroll.
[c]Although age ≥ 18 years, marked *prefer not to answer* for all items on the Expectancy Bias Task at screening.
[d]Superscript number reflects the number of these participants who did not start TX (i.e., were already non-ITT).
[e]Superscript number reflects the number of these participants who started but did not complete S4 (i.e., were already non-PP).