Empirical Article

Clinical Psychological Science
2023, Vol. 11(5) 819–840

DOI: 10.1177/21677026221103128
www.psychologicalscience.org/CPS

Shifting Episodic Prediction With Online Cognitive Bias Modification: A Randomized Controlled Trial

Jeremy W. Eberle[1], Mehdi Boukhechba[2], Jianhui Sun[3], Diheng Zhang[1], Daniel H. Funk[4], Laura E. Barnes[2,5], and Bethany A. Teachman[1]

[1]Department of Psychology, University of Virginia; [2]Department of Engineering Systems and Environment, University of Virginia; [3]Department of Computer Science, University of Virginia; [4]Sartography, Staunton, Virginia; and [5]School of Data Science, University of Virginia

Abstract

Negative future thinking pervades emotional disorders. This hybrid efficacy–effectiveness trial tested a four-session, scalable online cognitive-bias-modification program for training more positive episodic prediction. Nine hundred fifty-eight adults (73.3% female, 86.5% White, 83.4% from United States) were randomly assigned to positive conditions with ambiguous future scenarios that ended positively, 50/50 conditions that ended positively or negatively, or a control condition with neutral scenarios. As hypothesized (preregistration: https://osf.io/jrst6), positive-training participants improved more than control participants in negative expectancy bias ($d = -0.58$), positive expectancy bias ($d = 0.80$), and self-efficacy ($d = 0.29$). Positive training was also superior to 50/50 training for expectancy bias and optimism ($d = 0.31$). Training gains attenuated yet remained by 1-month follow-up. Unexpectedly, participants across conditions improved comparably in anxiety and depression symptoms and growth mindset. Targeting a transdiagnostic process with a scalable program may improve bias and outlook; however, further validation of outcome measures is required.

Keywords

cognitive bias modification, prospection, episodic prediction, expectancy bias, transdiagnostic

Received 8/21/20; Revision accepted 4/22/22

Prospection, the representation of future events (Gilbert & Wilson, 2007), is a central feature of mental life and may influence psychological health (Seligman et al., 2013). Healthy prospection has benefits such as improved emotion regulation (Pham & Taylor, 1999) and problem-solving abilities (e.g., Miloyan & Suddendorf, 2015), whereas negative prospection may be a transdiagnostic process that maintains anxiety and mood disorders (Roepke & Seligman, 2016). Early cognitive models posited maladaptive prospection in these disorders (Beck et al., 1987), including biases in *episodic prediction*, which refers to estimates of the likelihood of future autobiographical events and one's reaction to them (Szpunar et al., 2014). In the current study, we focus on shifting episodic prediction to be less negative and more positive and examine effects on multiple markers of psychological health (Becker et al., 2011; Kazdin, 2001), including symptoms (i.e., anxiety, depression) and positive-outlook outcomes (i.e., self-efficacy, growth mindset, optimism) tied to positive functioning and mental and physical well-being (Carver & Scheier, 2018; Dweck & Yeager, 2018; Maddux & Kleiman, 2018). To create a scalable intervention for episodic prediction, we developed an online variant of cognitive bias modification (CBM), a group of interventions that uses cognitive science principles to target disorder-relevant processing

Corresponding Author:
Jeremy W. Eberle, Department of Psychology, University of Virginia
Email: jwe4ec@virginia.edu

biases. These approaches typically require no professional contact and can be completed anywhere with Internet access, making them accessible to people who otherwise may face a host of barriers to accessing mental health care (Teachman, 2014).

In an initial efficacy trial of a brief online interpretation-bias program adapted to target episodic prediction directly, Namaky et al. (2021) found promising results for shifting future-thinking and positive-outlook outcomes in college students with relatively negative expectations of the future (although no condition effects on symptom reduction). The present study is a hybrid efficacy–effectiveness trial that tests the effectiveness and feasibility of this program when delivered to a large sample of adults around the world on our team's public research website called MindTrails. Although testing an online, self-guided intervention in the community with no payment for participation comes with challenges (e.g., need to minimize assessment burden and high levels of attrition common to public intervention websites), taking this deployment-focused approach with end users in natural settings permits evaluation of potential scalability and accelerates implementation.

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 & Schutte, 2017; Meevissen et al., 2011) or ordinary positive future events (Quoidbach et al., 2009). In addition, participants have simulated positive future events in response to cue words that were self-generated (Szpunar & Schacter, 2013) or provided (Boland et al., 2018). Participants have also been instructed to imagine positive outcomes or the process of achieving them (Taylor et al., 1998). Finally, in variations on the *ambiguous scenarios paradigm* (Mathews & Mackintosh, 2000), a widely used form of CBM 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 (e.g., "You ask a friend to look over some work you have done. They come back with some comments, which are *all very positive* [resolution is presented in italics throughout article]," presented auditorily in Blackwell & Holmes, 2010). Some authors have proposed that resolving this ambiguity involves the generation of future imagery (Lang et al., 2012; Murphy et al., 2017), a form of episodic simulation, which is one mode of prospection (Szpunar et al., 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 et al., 2017), and optimism (Murphy et al., 2015).

Targeting Episodic Prediction Directly

In Namaky et al.'s (2021) variation of CBM-I using the ambiguous scenarios paradigm, the program was designed to target episodic prediction directly. Rather than using present-tense scenarios, Namaky et al. used scenarios that described short- and long-term future events and used future tense in resolving the ambiguity (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 *success*."). They also solicited future predictions in postscenario comprehension questions (e.g., "Will your performance probably contribute to the team's success?"). They found that in college students with more negative expectancies (relative to the large student group that was 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 participants 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. (2021) were promising, they were limited to a relatively small college sample; were 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); for example, only 10% of users complete a second module at MoodGym, a popular self-help cognitive behavior therapy website (Batterham et al., 2008). In the present study, we build on Namaky et al. by testing a similar program in a larger, broader sample and including mid- and posttraining assessments and a slightly longer follow-up period (1 month) on a platform easily disseminated 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 CBM intervention for reducing negative expectancy bias and increasing positive expectancy bias in community adults with relatively negative expectancies. Participants were randomly assigned to one of five conditions. The *positive prospection* condition was designed to train more positive episodic prediction via repeated practice envisioning positive outcomes to emotionally ambiguous, self-relevant future situations (see soccer league example from Namaky et al., 2021, above). A second positive condition, *positive prospection + negation*, supplements the envisioning of positive outcomes with a negation of negative outcomes (e.g., "You believe that you will *not let your teammates down, and* contribute to your team's succ_ss."; emphasis added) to test the impact of adding negation, given 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). For each of these positive conditions, although most (90%) of the scenarios end positively, some (10%) of the scenarios end negatively (e.g., "fai_ure") to reduce responding on "autopilot" and to retain some uncertainty about the outcomes as the scenarios unfold. This was done given the importance of (a) emotional ambiguity for shifting interpretation bias (Clarke et al., 2014) and (b) uncertainty and surprise in belief adjustment following prediction error (McGuire et al., 2014).

Whereas the two positive conditions were designed to train a positive contingency while retaining some flexibility about the possibility of a negative outcome, two 50/50 conditions were designed to emphasize flexibility. Presenting equal proportions of positive (P; 50%) and negative (N; 50%) outcomes, the 50/50 conditions provide repeated practice envisioning different outcomes to emotionally ambiguous future situations without training a contingency. The *50/50 random* condition uses a random order of valence (e.g., PNNPNPPNPN), whereas the *50/50 blocked* condition uses five-scenario blocks of alternating valence (e.g., PPPPPNNNNN), which may promote greater flexibility by requiring that participants shift their future thinking after developing a pattern of positive or negative expectation in the prior block. Finally, the *neutral control* condition controls for the CBM format of the other conditions but uses situations that lack emotional ambiguity about the future and end 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."). Thus, whereas the positive conditions train a positive contingency and the positive conditions and 50/50 conditions both resolve emotional ambiguity, the neutral control task does not train a contingency or resolve emotional ambiguity.

We preregistered several directional hypotheses (https://osf.io/jrst6). First, during training, 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 participants in the neutral control condition. Second, participants in the two 50/50 conditions will improve significantly more than participants in the neutral control condition but less than participants in the two positive conditions. Third, participants in the 50/50 blocked condition will improve significantly more than participants in the 50/50 random condition. Finally, we preregistered a nondirectional test of differential improvements between the two positive conditions.

Method

Participants and design

Enrollment and data collection began on May 3, 2017; enrollment ended on January 20, 2019; and data collection ended on October 16, 2019. Following our preregistration (https://osf.io/jrst6), we analyzed data for participants who enrolled on or before March 27, 2018. We analyzed data collected for these participants through September 9, 2018. A sample of 4,751 community participants self-selected to complete a screening on the MindTrails Project website (https://mindtrails.virginia.edu) "to encourage healthier thinking about the future for people who tend to expect things will not turn out well." The web-based program could be completed on computers, tablets, and smartphones. In all, 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 randomly assigned to one of five conditions (for details, see CONSORT diagram in Fig. 1 and Section S1.1 in Supplemental Material available online). Procedures for all participants, who were not told their condition, were identical from the point of random assignment to the start of the first training session. Participants were encouraged to complete a pretraining assessment within 2 days. After they did, they were immediately given the opportunity to begin the first training session. Nine hundred seventy-one participants began the first session (i.e., viewed at least the first training scenario); 13 of these participants (one who had submitted a

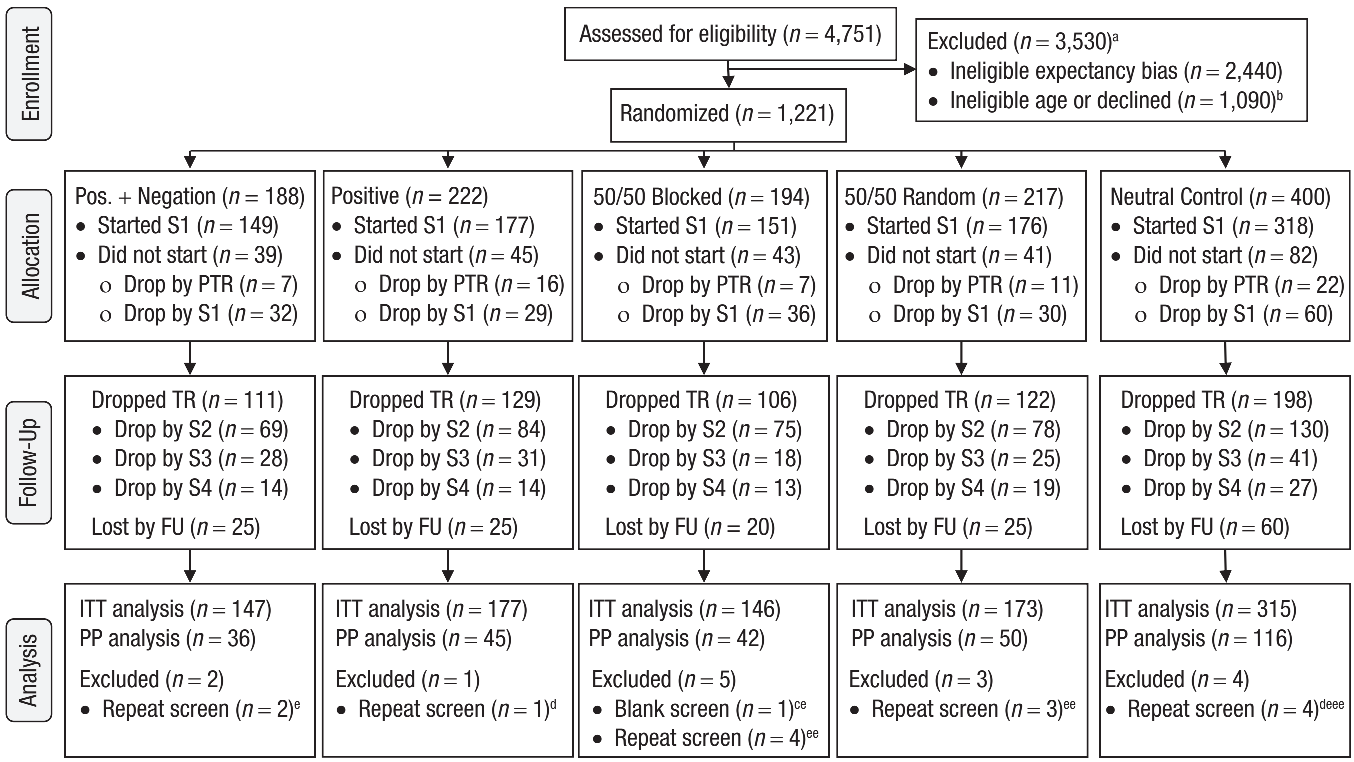

Fig. 1. Participant flowchart. Numbers dropped by a given session reflect participants who did not start (vs. started but did not complete) the session. Numbers lost by follow-up reflect participants who did not start (vs. started but did not complete) follow-up. S1–S4 = Session 1–Session 4; PTR = pretraining; TR = training; 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 S1 (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).

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. Two hundred eighty-nine of these ITT participants completed all four training sessions and form the per-protocol (PP) sample. The University of Virginia Institutional Review Board approved all 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)[1] 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-time (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 participants 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 [PHQ-4], see below). For full demographic information, see Table S2 in the Supplemental Material, and for current and lifetime diagnoses (which participants self-reported receiving from a health professional) and ancillary treatment or social support for mental or emotional difficulties, see Table S3 in the Supplemental Material.

Recruitment. The MindTrails website and present study were advertised through university online press releases, local and statewide news (e.g., radio interviews), Craigslist postings, emails to clinicians, flyers at university counseling centers, a link at the Project Implicit Health website, and the MindTrails Project Facebook page. A link to the original MindTrails study (Ji et al., 2021) was posted on ClinicalTrials.gov (Identifier NCT02382003).

Training and assessment schedule. Participants were asked to complete four training sessions (two per week, 2–4 days apart), following Namaky et al. (2021). This number of sessions was chosen because (a) CBM-I studies with multiple sessions have shown larger effects than studies with one session (Menne-Lothmann et al., 2014) and (b) other studies based on the ambiguous scenarios paradigm have shown effects with four or fewer sessions (e.g., Bowler et al., 2012; Namaky et al., 2021). Assessments were given in a fixed order immediately after each session and at 1-month follow-up (because training and assessment occurred in an 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 in the Supplemental Material).

Measures

Expectancy bias. Expectancy bias was assessed with a modified Expectancy Bias Task (Namaky et al., 2021), 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, in which experiences seldom consist of only positive or only negative events, and the four scenarios described four of the six domains targeted in training—health, family/friends, evaluations/performance, and finances—without overlapping in content with training 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 through 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., 2021). 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. For more details, see Section S1.4 in the Supplemental Material.

To understand the effects of the intervention 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 (ω_t) using complete item-level data[2] at pretraining was unacceptable for the ITT sample for the negative events (ω_t = .28, 95% confidence interval [CI] [.14, .35]) and the positive events (ω_t = .31, 95% CI [.23, .38]), and plausible estimates and stable standard errors for the PP sample did not emerge.[3] We had assumed one dimension for the negative events and one for the positive events, but confirmatory factor analyses for each latent factor with the *OpenMx* package (Version 2.17.3; Neale et al., 2016) in R (R Core Team, 2020) showed poor model fit for the ITT and PP samples (see Section S1.6 and Table S5 in the Supplemental Material), which is a limitation of this measure.

Anxiety and depression symptoms. Anxiety and depression symptoms were assessed with the PHQ-4 (Kroenke et al., 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 anxiety items, identical to those of the Generalized Anxiety Disorder–2 (GAD-2; Kroenke et al., 2007) scale, 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 et al., 2003), comprise the Depression subscale, for which a sum of 3 or greater reflects potential major depression or another depressive disorder. Prior studies support the construct and criterion validity of the GAD-2 and PHQ-2 (Kroenke et al., 2003, 2007, 2009; Löwe et al., 2005), and their internal consistency, discriminative validity, 1- to 4-week test-retest reliability, and sensitivity to treatment change have been found comparable to those of longer measures (Staples et al., 2019; see also Kroenke et al., 2010). In the present study, the PHQ-4 was administered at pretraining, 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 pretraining was good for the anxiety items (ITT: ω_t = .82, 95% CI [.79, .84]; PP: ω_t = .83, 95% CI [.77, .87]) and the depression items (ITT: ω_t = .81, 95% CI [.78, .84]; PP: ω_t = .83, 95% CI [.77, .87]).

Self-efficacy, growth mindset, and optimism. Self-efficacy, growth mindset, and optimism were assessed at pretraining, 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., 2021), we identified two or three items per scale using data from Namaky et al. (2021). To improve user experience, we made some small modifications to wording (noted below) and used a consistent response format across items—a 5-point Likert scale ranging from 0 (*strongly disagree*) to 4 (*strongly agree*). For details, see Section S1.7 in the Supplemental Material. 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 from original); "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 et al., 2001). Internal consistency for the ITT and PP samples using complete item-level data at pretraining was good (ITT: ω_t = .82, 95% CI [.79, .84]; PP: ω_t = .83, 95% CI [.79, .86]).

Growth mindset was assessed with three items from a set of growth-mindset questions (GMQ) about intelligence (Dweck, 2006). The selected items were adapted to make the target changing one's thinking rather than one's intelligence: "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 pretraining was good (ITT: ω_t = .80, 95% CI [.78, .83]; PP: ω_t = .83, 95% CI [.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 from original) and "I hardly ever expect things to go my way" (both reverse scored; Scheier et al., 1994). Internal consistency for the ITT and PP samples using complete item-level data at pretraining was good (ITT: ω_t = .80, 95% CI [.77, .83]; PP: ω_t = .81, 95% CI [.75, .85]).

Training confidence and change importance. Training confidence and change importance were assessed at pretraining, before any exposure to different training conditions. Training confidence was assessed with two 5-point Likert items, modified from Borkovec and Nau (1972), and ranging from 0 (*not at all*) to 4 (*extremely*): "How confident are you that an online training program that is designed to change how you think about situations will be successful in changing your thinking about your future?" and "How confident are you that an online training program will be successful in changing your thinking about your future?" Change importance was assessed on the same scale with a single item modified from the Importance Ruler (Center for Evidence-Based Practices, 2010): "How important is changing your thinking about your future to you right now?" Given some evidence that the confidence in CBM-I training predicts outcomes and dropout (Hohensee et al., 2020), this item and the mean of the training-confidence items were analyzed along with demographic variables as potential predictors of dropout (see Missing Data Handling section below and Section 1.12 in the Supplemental Material).

Conditions

At the start of each session, participants completed two questions about current positive and negative feelings and then their condition's tasks. Each of the five conditions' tasks 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 (for details, see Sections S1.8 and S1.9 in the Supplemental Material).

CBM tasks. In the CBM tasks, the 40 scenarios in each session were randomly chosen from a set of 49. Scenarios were not repeated in 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 the positive prospection condition, 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, and 10% ended negatively. In 50/50 blocked, 50% of the scenarios ended positively, and 50% ended negatively; the valence alternated 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 and 2, only one letter was missing (e.g., see above); in Sessions 3 and 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 and 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 in 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. As in the CBM tasks, the number of missing letters varied across sessions (one in Sessions 1 and 2, two in Sessions 3 and 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; each additional sentence after the first appeared after 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 article are two-tailed, and the alpha level is .05 (except in multilevel models, in which the Bonferroni-corrected level is .025—see the section Multilevel Modeling). For initial tests of baseline demographic differences, see Section S1.10 in the Supplemental Material, and for descriptions of deviations from our preregistered analytic plan, see Section S2 in the Supplemental Material.

Longitudinal outcomes.

Preprocessing. To compare longitudinal outcomes between multiple combinations of conditions, two versions of the data set with 958 ITT participants were created: a combined-level data set with the condition variable dummy coded in three levels—(a) both positive conditions (including the positive prospection and positive prospection + negation conditions), (b) both 50/50 conditions (including the 50/50 blocked and 50/50 random conditions), and (c) the neutral control condition—and a separate-level data set with the condition variable dummy-coded in five levels, one for each of the five conditions (see Section S.1.11 in the Supplemental Material). Dummy rather than contrast coding was used because in the combined-level data set, it 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 the both positive level).

Missing data handling. The present data exhibit 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% to 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% (for the number of observations of each outcome over time per condition, see Table S1 in the Supplemental Material). To identify measured variables other than time that may relate to this pattern of missing data, we tested whether training confidence, change importance, and demographic variables predicted the number of missing sessions (using nonparametric tests because this number is not normally distributed, but negatively skewed). Age and education were the only significant predictors (for details, see Section S1.12 in the Supplemental Material).

Following an inclusive analysis strategy (Collins et al., 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 data sets, we used the *jomo* (Version 2.7-2; Quartagno & Carpenter, 2020) and *mitml* packages (Version 0.4-1; Grund et al., 2021) in R (R Core Team, 2021) 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 (assessment point), which we represented with two linear splines in the same model: time_{TR} for the training trajectory (coded as 0 for baseline, as integers 1–4 for Sessions 1–4, and as 4 for follow-up) and time_{FU} for the follow-up trajectory (coded as 0 for baseline–Session 4 and as 1 for follow-up). The Level 1 predictors were the fixed effects of condition, time_{TR}, time_{FU}, Condition × Time_{TR} interaction, and Condition × Time_{FU} interaction and random effects for intercept and time_{TR}. The Level 2 predictor was condition.[4] We treated target variables as continuous and followed Grund et al. (2018) to specify the model and impute 20 data sets (see Section S1.14 in the Supplemental Material). Multivariate normal and missing-at-random data were assumed for the imputation and subsequent analysis models. Because the PHQ-4, NGSE, GMQ, and LOT-R were not assessed at Sessions 1 or 3, imputed data at these time points were removed before analysis of anxiety, depression, self-efficacy, growth mindset, and optimism outcomes.

Multilevel modeling. We conducted analyses separately for each of the imputed data sets, and following Rubin's rules, we 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 *confint* function was used to compute Bonferroni-corrected 97.5% CIs (see below) for final estimates, all of which we report in terms of unstandardized *b*.

Differential change over time between conditions was assessed using hierarchical linear models with restricted maximum likelihood estimation specified using the *nlme* package (Version 3.1-152; Pinheiro et al., 2021) in R (R Core Team, 2021). We used the *optim* optimizer (Broyden–Fletcher–Goldfarb–Shanno method), which reduced convergence errors obtained with the default *nlminb* optimizer. We assumed piecewise linear trajectories (one during training, one during follow-up) and in each model simultaneously entered fixed effects of condition, time_{TR}, time_{FU}, and the Condition × Time_{TR} and Condition × Time_{FU} interactions and random effects for intercept and time_{TR} (because the follow-up phase had only two time points, a random slope was not entered for time_{FU}).

The combined-level data set was used to compare (a) the two positive conditions with the neutral control condition, (b) the two 50/50 conditions with the neutral control condition, and (c) the two positive conditions with the two 50/50 conditions. The separate-level data set 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 data set) or the lower-order fixed effects. Given that only two of the interactions we interpreted per data set 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 fixed effects of time_{TR} and time_{FU} and random effects for intercept and time_{TR} (for ITT and PP samples, see Table 2 and Table S14 in the Supplemental Material). Finally, because in the models containing condition, time_{TR}, time_{FU}, and the Condition × Time_{TR} and Condition × Time_{FU} interactions the estimates for the main effects of time_{TR} and time_{FU} vary according to which condition is specified as the reference group for the dummy condition variable, the simple effects of time_{TR} and time_{FU} in all five conditions were assessed in the separate-level data set to understand the overall change in all conditions regardless of the interactions' significance (for ITT and PP samples, see Table 1 and Table S13 in the Supplemental Material). We did this instead of testing the main effects of time_{TR} and time_{FU} across all five conditions because main effects are misleading when interactions are 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*. GMA *d* was computed at the end of the training phase (i.e., Session 4) and at the end of the follow-up phase (i.e., follow-up) using the pooled within-groups standard deviation at baseline. For each case, we adapted Feingold's (2018) Equation 2 for time-varying effect sizes for quadratic growth models to linear spline models: GMA $d = (b_{TR} \times \text{time}_{TR} + b_{FU} \times \text{time}_{FU}) / SD$, where b_{TR} is the coefficient for the Condition × Time_{TR} interaction effect, b_{FU} is the coefficient for the Condition × Time_{FU} interaction effect, and time_{TR} and time_{FU} are the codings for these variables at the desired time point. Thus, for GMA d at end of training, we computed $(b_{TR} \times 4 + b_{FU} \times 0) / SD$, and for GMA d at follow-up, we computed $(b_{TR} \times 4 + b_{FU} \times 1) / SD$. Given that post hoc equations for computing CIs for time-varying GMA *d*s for quadratic (or linear spline) models have not been derived, we did not compute CIs for GMA *d*s.

For the within-group effect size of the interpreted simple effects of time, we summed the product of the coefficient for the time_{TR} slope and the coding of time_{TR} at the desired time point (i.e., Session 4, follow-up) and the product of the coefficient for the time_{FU} slope and the coding of time_{FU} at that time point (yielding a numerator equal to the difference between the group's estimated means at baseline and at the desired time point) and divided by the group's standard deviation at baseline. This yields a GMA analogue of effect size for a one-group pretest-posttest design; CIs were not computed for these effect sizes because equations for the standard errors have not been derived (A. Feingold, personal communication, March 3–4, 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 percentage decrease on this score from screening to each assessment point on the basis of 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 percentage decrease (for details, see Section S1.15 in the Supplemental Material). 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

Because the ITT analyses retain groups that systematically differ only by randomized condition, permitting causal inferences about training effects, we report only ITT results below (see Altman, 2009; Hollis & Campbell, 1999). PP analyses revealed similar results, but with fewer significant effects, presumably because of reduced power given the smaller sample and to potentially biased estimates given nonrandom attrition in the full ITT sample (for full PP results, see Tables S13 and S14

Table 1. Linear Spline Multilevel Modeling Time Effects in Each Condition for Intent-to-Treat Sample

Outcome	Phase	Condition	*b* (*SE*)	*df*	*t*	*p*	*d*[a]
Positive bias	TR	Positive + negation	0.51 (0.04)	51.34	12.72	< .001**	2.08
		Positive	0.54 (0.03)	57.09	15.83	< .001**	2.27
		50/50 blocked	0.34 (0.04)	44.09	8.44	< .001**	1.56
		50/50 random	0.43 (0.04)	64.24	12.19	< .001**	1.73
		Neutral control	0.34 (0.03)	49.13	12.20	< .001**	1.34
	FU	Positive + negation	−0.45 (0.26)	16.25	−1.71	.106	1.62
		Positive	−0.63 (0.19)	21.18	−3.40	.003**	1.60
		50/50 blocked	−0.26 (0.17)	23.53	−1.59	.126	1.26
		50/50 random	−0.18 (0.18)	21.68	−0.99	.333	1.55
		Neutral control	−0.36 (0.12)	24.03	−2.97	.007*	0.98
Negative bias	TR	Positive + negation	−0.34 (0.04)	43.76	−9.16	< .001**	−1.33
		Positive	−0.36 (0.03)	56.34	−11.89	< .001**	−1.60
		50/50 blocked	−0.21 (0.04)	40.92	−5.53	< .001**	−0.78
		50/50 random	−0.21 (0.03)	53.91	−6.78	< .001**	−0.92
		Neutral control	−0.22 (0.02)	72.12	−10.05	< .001**	−0.91
	FU	Positive + negation	0.38 (0.22)	17.23	1.70	.107	−0.95
		Positive	0.36 (0.20)	18.02	1.76	.095	−1.20
		50/50 blocked	0.08 (0.18)	21.01	0.47	.640	−0.70
		50/50 random	−0.24 (0.17)	21.12	−1.42	.171	−1.18
		Neutral control	0.11 (0.13)	21.70	0.87	.393	−0.79

(continued)

Table 1. *(continued)*

Outcome	Phase	Condition	*b* (*SE*)	*df*	*t*	*p*	*d*[a]
Anxiety	TR	Positive + negation	−0.32 (0.07)	33.96	−4.84	< .001**	−0.71
		Positive	−0.17 (0.06)	40.79	−2.96	.005*	−0.34
		50/50 blocked	−0.19 (0.06)	41.76	−3.26	.002**	−0.43
		50/50 random	−0.14 (0.05)	54.34	−2.61	.012*	−0.28
		Neutral control	−0.21 (0.04)	51.76	−5.31	< .001**	−0.45
	FU	Positive + negation	0.42 (0.46)	12.65	0.92	.375	−0.47
		Positive	−0.01 (0.29)	21.19	−0.04	.971	−0.35
		50/50 blocked	0.11 (0.27)	24.42	0.40	.696	−0.37
		50/50 random	−0.80 (0.29)	21.44	−2.81	.010*	−0.69
		Neutral control	0.11 (0.18)	28.49	0.60	.551	−0.39
Depression	TR	Positive + negation	−0.22 (0.07)	33.15	−3.38	.002**	−0.46
		Positive	−0.19 (0.06)	36.40	−3.54	.001**	−0.40
		50/50 blocked	−0.13 (0.07)	30.58	−1.97	.058	−0.30
		50/50 random	−0.26 (0.06)	39.38	−4.44	< .001**	−0.51
		Neutral control	−0.21 (0.04)	63.46	−5.83	< .001**	−0.44
	FU	Positive + negation	0.36 (0.35)	16.84	1.04	.312	−0.27
		Positive	0.09 (0.29)	19.47	0.30	.769	−0.36
		50/50 blocked	−0.26 (0.32)	19.00	−0.81	.428	−0.45
		50/50 random	0.09 (0.28)	22.30	0.33	.745	−0.47
		Neutral control	0.26 (0.17)	28.52	1.55	.131	−0.30
Self-efficacy	TR	Positive + negation	0.16 (0.03)	29.04	5.21	< .001**	0.77
		Positive	0.20 (0.03)	33.28	7.13	< .001**	0.91
		50/50 blocked	0.13 (0.03)	34.06	4.26	< .001**	0.58
		50/50 random	0.16 (0.02)	90.37	6.94	< .001**	0.69
		Neutral control	0.12 (0.02)	67.28	7.31	< .001**	0.52
	FU	Positive + negation	−0.31 (0.17)	14.40	−1.84	.086	0.41
		Positive	−0.14 (0.11)	22.15	−1.24	.227	0.75
		50/50 blocked	0.08 (0.13)	20.16	0.67	.514	0.67
		50/50 random	−0.19 (0.11)	24.55	−1.75	.092	0.49
		Neutral control	−0.02 (0.09)	22.47	−0.27	.788	0.50
Growth mindset	TR	Positive + negation	0.14 (0.03)	31.77	3.97	< .001**	0.66
		Positive	0.11 (0.03)	36.57	4.00	< .001**	0.48
		50/50 blocked	0.11 (0.03)	42.52	4.01	< .001**	0.54
		50/50 random	0.15 (0.03)	34.95	4.91	< .001**	0.61
		Neutral control	0.08 (0.02)	41.77	3.85	< .001**	0.36
	FU	Positive + negation	0.07 (0.17)	15.37	0.41	.685	0.74
		Positive	0.09 (0.14)	17.34	0.65	.523	0.58
		50/50 blocked	0.16 (0.11)	25.24	1.47	.155	0.75
		50/50 random	−0.15 (0.11)	24.25	−1.34	.194	0.45
		Neutral control	−0.02 (0.07)	29.07	−0.30	.765	0.33
Optimism	TR	Positive + negation	0.19 (0.04)	26.84	5.46	< .001**	0.79
		Positive	0.14 (0.03)	33.67	4.88	< .001**	0.62
		50/50 blocked	0.12 (0.03)	45.46	4.26	< .001**	0.49
		50/50 random	0.09 (0.02)	68.89	3.77	< .001**	0.37
		Neutral control	0.11 (0.02)	48.16	5.71	< .001**	0.45
	FU	Positive + negation	−0.02 (0.18)	13.68	−0.10	.920	0.77
		Positive	0.14 (0.12)	19.99	1.16	.261	0.78
		50/50 blocked	0.26 (0.13)	20.62	2.02	.056	0.75
		50/50 random	0.28 (0.12)	20.58	2.35	.029	0.66
		Neutral control	−0.10 (0.09)	21.88	−1.13	.273	0.34

Note: Separate models were fit for each outcome and condition. Every model included the fixed effects of time_{TR} (TR phase trajectory) and time_{FU} (FU phase trajectory), a random intercept, and a random slope for time_{TR}. The separate-level data set, 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. TR = training; FU = follow-up.
[a]For TR phase, *d* is the standardized mean difference in a given condition from baseline to Session 4. For FU phase, *d* is the standardized mean difference from baseline to FU and is computed from both the time_{TR} and time_{FU} effects.
*$p < .025$. **$p < .005$.

in the Supplemental Material). Moreover, no significant slope difference emerged during training 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 sample (for full results for these comparisons for the ITT and PP samples, see Table 2 and Table S14 in the Supplemental material). Therefore, we focus below on differential change for the combined positive and combined 50/50 conditions.

Positive expectancy bias. Regarding within-groups change, ITT participants in all five conditions significantly improved in positive bias during training (*b*s = 0.34–0.54, *p*s < .001), with pre-post effect sizes *d* of 1.34 to 2.27 at Session 4 (Table 1). Participants in most individual conditions showed no significant change from Session 4 to follow-up. Although participants in the positive condition and neutral control condition showed significant losses in training gains (*b*s = −0.63 and −0.36, largest *p* = .003), within-groups effect sizes from baseline to follow-up remained positive (*d*s = 1.60 and 0.98), suggesting that these participants still improved overall. Regarding between-groups change, as hypothesized, during training participants in the two positive conditions improved significantly more than participants in the neutral control condition (*b* = 0.20, *p* < .001, *d* = 0.80) and participants in the two 50/50 conditions (*b* = 0.15, *p* < .001, *d* = 0.63; Table 2; Fig. 2). However, no significant slope difference emerged between the two 50/50 conditions and the neutral control condition. Although the positive conditions significantly decreased in positive bias relative to the 50/50 conditions from Session 4 to follow-up (*b* = −0.34, *p* = .011), the between-groups effect size from baseline to follow-up still favored the positive conditions overall (*d* = 0.28). No significant slope difference emerged between the positive conditions and the neutral control condition from Session 4 to follow-up, suggesting maintenance of the positive conditions' superiority with a between-groups effect size *d* from baseline to follow-up of 0.61.

Negative expectancy bias. Showing findings similar to those for positive bias, ITT participants in all five conditions significantly improved in negative bias during training (*b*s = −0.21 to −0.36, *p*s < .001, *d*s = −0.78 to −1.60; Table 1). Participants in every individual condition showed no significant change from Session 4 to follow-up. As expected, during training, participants in the two positive conditions improved significantly more than participants in the neutral control condition (*b* = −0.14, *p* < .001, *d* = −0.58) and participants in the two 50/50 conditions (*b* = −0.16, *p* < .001, *d* = −0.64; Table 2; Fig. 2). Again, no significant slope difference emerged between the two 50/50 conditions and the neutral control condition. Although the positive conditions together showed a significant loss in training gains from Session 4 to follow-up (*b* = 0.41, *p* = .001), significantly increasing in negative bias relative to the 50/50 conditions (*b* = 0.50, *p* = .005), the positive conditions' within-group effect size from baseline to follow-up remained negative (*d* = −1.06), suggesting improvement overall, and the between-groups effect size from baseline to follow-up favored the positive conditions (*d* = −0.13). No significant slope difference emerged between the positive conditions and the neutral control condition from Session 4 to follow-up, suggesting maintenance of the positive conditions' superiority with a between-groups effect size *d* from baseline to follow-up of −0.27.

Anxiety symptoms. ITT participants in all five conditions significantly improved in anxiety symptoms during training (*b*s = −0.14 to −0.32, largest *p* = .012, *d*s = −0.28 to −0.71; Table 1). Participants in 50/50 random continued to significantly improve from Session 4 to follow-up (*b* = −0.80, *p* = .010), with a within-group effect size *d* from baseline to follow-up of −0.69; no other condition significantly changed from Session 4 to follow-up. Contrary to our hypotheses, no significant slope difference emerged between the conditions we compared during training or from Session 4 to follow-up (Table 2; see Fig. S8 in the Supplemental Material).

Depression symptoms. Showing findings similar to those for anxiety symptoms, ITT participants in all five conditions except 50/50 blocked significantly improved in depression symptoms during training (*b*s = −0.19 to −0.26, largest *p* = .002, *d*s = −0.40 to −0.51; Table 1). No condition significantly changed from Session 4 to follow-up, and no significant slope difference emerged between the conditions we compared during training or from Session 4 to follow-up (Table 2; see Fig. S8 in the Supplemental Material).

Self-efficacy. ITT participants in all five conditions significantly improved in self-efficacy during training (*b*s = 0.12–0.20, *p*s < .001, *d*s = 0.52–0.91), and no condition significantly changed from Session 4 to follow-up (Table 1). As hypothesized, during training, participants in the two positive conditions improved significantly more than participants in the neutral control condition (*b* = 0.07, *p* = .009, *d* = 0.29), with no significant slope difference between the positive conditions and the neutral control condition from Session 4 to follow-up, yielding a between-groups effect size *d* from baseline to follow-up of 0.11 (Table 2; Fig. 2). No significant slope difference emerged between other conditions we compared during training or from Session 4 to follow-up.

Table 2. Linear Spline Multilevel Modeling Fixed Condition × Time Interaction and Simple Time Effects for Intent-to-Treat Sample

Outcome	Phase	Effect	*b* (*SE*)	97.5% CI	*df*	*t*	*p*	*d*[a]
Positive bias	TR	(Both Positive vs. Neutral Control) × Time	0.20 (0.04)	[0.11, 0.28]	64.16	5.36	< .001**	0.80
		$\text{Time}_{\text{Both Positive}}$	0.53 (0.03)	[0.46, 0.60]	39.62	17.27	< .001**	2.21
		$\text{Time}_{\text{Neutral Control}}$	0.33 (0.03)	[0.27, 0.40]	49.13	12.20	< .001**	1.34
		(Both Positive vs. Both 50/50) × Time	0.15 (0.04)	[0.06, 0.24]	55.48	3.90	< .001**	0.63
		$\text{Time}_{\text{Both Positive}}$	0.53 (0.03)	[0.46, 0.60]	39.62	17.27	< .001**	2.21
		$\text{Time}_{\text{Both 50/50}}$	0.38 (0.03)	[0.31, 0.46]	38.63	12.26	< .001**	1.64
		(Both 50/50 vs. Neutral Control) × Time	0.05 (0.03)	[−0.03, 0.12]	107.88	1.44	.153	0.19
		(Positive + Negation vs. Positive) × Time	−0.02 (0.05)	[−0.14, 0.09]	68.65	−0.47	.639	−0.10
		(50/50 Blocked vs. 50/50 Random) × Time	−0.09 (0.05)	[−0.20, 0.03]	65.23	−1.68	.098	−0.37
	FU	(Both Positive vs. Neutral Control) × Time	−0.18 (0.15)	[−0.54, 0.18]	29.61	−1.20	.240	0.61
		(Both Positive vs. Both 50/50) × Time	−0.34 (0.13)	[−0.63, −0.04]	38.39	−2.67	.011*	0.28
		$\text{Time}_{\text{Both Positive}}$	−0.54 (0.11)	[−0.79, −0.29]	29.35	−5.15	< .001**	1.65
		$\text{Time}_{\text{Both 50/50}}$	−0.21 (0.11)	[−0.48, 0.07]	26.76	−1.80	.083	1.42
		(Both 50/50 vs. Neutral Control) × Time	0.15 (0.16)	[−0.23, 0.54]	27.81	0.95	.352	0.35
		(Positive + Negation vs. Positive) × Time	0.18 (0.31)	[−0.56, 0.91]	22.58	0.57	.573	0.08
		(50/50 Blocked vs. 50/50 Random) × Time	−0.08 (0.21)	[−0.57, 0.41]	30.96	−0.39	.699	−0.46
Negative bias	TR	(Both Positive vs. Neutral Control) × Time	−0.14 (0.03)	[−0.21, −0.06]	60.25	−4.25	< .001**	−0.58
		$\text{Time}_{\text{Both Positive}}$	−0.36 (0.02)	[−0.41, −0.30]	55.86	−15.60	< .001**	−1.48
		$\text{Time}_{\text{Neutral Control}}$	−0.22 (0.02)	[−0.27, −0.17]	72.12	−10.05	< .001**	−0.91
		(Both Positive vs. Both 50/50) × Time	−0.16 (0.04)	[−0.24, −0.08]	49.98	−4.50	< .001**	−0.64
		$\text{Time}_{\text{Both Positive}}$	−0.36 (0.02)	[−0.41, −0.30]	55.86	−15.60	< .001**	−1.48
		$\text{Time}_{\text{Both 50/50}}$	−0.20 (0.03)	[−0.26, −0.14]	45.89	−7.90	< .001**	−0.81
		(Both 50/50 vs. Neutral Control) × Time	0.02 (0.03)	[−0.05, 0.09]	78.70	0.58	.567	0.07
		(Positive + Negation vs. Positive) × Time	0.02 (0.05)	[−0.08, 0.13]	62.84	0.48	.636	0.09
		(50/50 Blocked vs. 50/50 Random) × Time	0.00 (0.04)	[−0.10, 0.10]	72.97	−0.02	.987	0.00
	FU	(Both Positive vs. Neutral Control) × Time	0.30 (0.15)	[−0.07, 0.66]	27.22	1.95	.062	−0.27
		(Both Positive vs. Both 50/50) × Time	0.50 (0.16)	[0.11, 0.88]	25.83	3.06	.005*	−0.13
		$\text{Time}_{\text{Both Positive}}$	0.41 (0.11)	[0.15, 0.67]	25.67	3.76	.001**	−1.06
		$\text{Time}_{\text{Both 50/50}}$	−0.09 (0.12)	[−0.38, 0.20]	23.59	−0.72	.479	−0.90
		(Both 50/50 vs. Neutral Control) × Time	−0.20 (0.15)	[−0.55, 0.16]	27.92	−1.33	.195	−0.13
		(Positive + Negation vs. Positive) × Time	0.02 (0.26)	[−0.61, 0.65]	23.42	0.09	.933	0.11
		(50/50 Blocked vs. 50/50 Random) × Time	0.32 (0.25)	[−0.27, 0.92]	24.33	1.30	.207	0.32
Anxiety	TR	(Both Positive vs. Neutral Control) × Time	−0.02 (0.05)	[−0.14, 0.10]	73.24	−0.37	.716	−0.04
		(Both Positive vs. Both 50/50) × Time	−0.09 (0.06)	[−0.22, 0.04]	58.38	−1.65	.104	−0.19
		(Both 50/50 vs. Neutral Control) × Time	0.07 (0.05)	[−0.05, 0.20]	61.12	1.32	.192	0.15
		(Positive + Negation vs. Positive) × Time	−0.15 (0.08)	[−0.34, 0.04]	51.09	−1.85	.071	−0.32
		(50/50 Blocked vs. 50/50 Random) × Time	−0.05 (0.09)	[−0.25, 0.14]	46.86	−0.63	.531	−0.11
	FU	(Both Positive vs. Neutral Control) × Time	0.08 (0.28)	[−0.59, 0.74]	27.78	0.28	.783	0.00
		(Both Positive vs. Both 50/50) × Time	0.53 (0.29)	[−0.15, 1.20]	27.34	1.85	.075	0.09
		(Both 50/50 vs. Neutral Control) × Time	−0.45 (0.26)	[−1.07, 0.17]	30.19	−1.72	.097	−0.09
		(Positive + Negation vs. Positive) × Time	0.44 (0.59)	[−0.99, 1.86]	20.87	0.74	.471	−0.09
		(50/50 Blocked vs. 50/50 Random) × Time	0.91 (0.42)	[−0.10, 1.92]	26.22	2.15	.041	0.38

(continued)

Table 2. *(continued)*

Outcome	Phase	Effect	*b* (*SE*)	97.5% CI	*df*	*t*	*p*	d^a
Depression	TR	(Both Positive vs. Neutral Control) × Time	0.02 (0.05)	[–0.10, 0.13]	70.37	0.30	.769	0.03
		(Both Positive vs. Both 50/50) × Time	–0.01 (0.05)	[–0.13, 0.11]	58.74	–0.20	.840	–0.02
		(Both 50/50 vs. Neutral Control) × Time	0.03 (0.05)	[–0.09, 0.14]	67.33	0.50	.619	0.05
		(Positive + Negation vs. Positive) × Time	–0.03 (0.08)	[–0.20, 0.15]	58.46	–0.34	.735	–0.05
		(50/50 Blocked vs. 50/50 Random) × Time	0.13 (0.09)	[–0.07, 0.32]	43.60	1.50	.141	0.27
	FU	(Both Positive vs. Neutral Control) × Time	–0.13 (0.22)	[–0.65, 0.40]	35.00	–0.56	.581	–0.03
		(Both Positive vs. Both 50/50) × Time	0.21 (0.26)	[–0.41, 0.83]	28.60	0.81	.426	0.09
		(Both 50/50 vs. Neutral Control) × Time	–0.34 (0.27)	[–0.97, 0.30]	27.95	–1.25	.221	–0.12
		(Positive + Negation vs. Positive) × Time	0.28 (0.43)	[–0.74, 1.30]	25.00	0.65	.523	0.09
		(50/50 Blocked vs. 50/50 Random) × Time	–0.35 (0.41)	[–1.33, 0.63]	26.06	–0.85	.402	0.09
Self-efficacy	TR	(Both Positive vs. Neutral Control) × Time	0.07 (0.02)	[0.01, 0.12]	69.18	2.68	.009*	0.29
		$\text{Time}_{\text{Both Positive}}$	0.19 (0.02)	[0.14, 0.23]	42.96	9.71	< .001**	0.85
		$\text{Time}_{\text{Neutral Control}}$	0.12 (0.02)	[0.08, 0.16]	67.28	7.31	< .001**	0.52
		(Both Positive vs. Both 50/50) × Time	0.04 (0.03)	[–0.02, 0.10]	50.57	1.48	.144	0.18
		(Both 50/50 vs. Neutral Control) × Time	0.03 (0.02)	[–0.03, 0.08]	87.70	1.10	.276	0.11
		(Positive + Negation vs. Positive) × Time	–0.04 (0.04)	[–0.14, 0.06]	39.02	–0.84	.406	–0.17
		(50/50 Blocked vs. 50/50 Random) × Time	–0.03 (0.04)	[–0.11, 0.05]	65.76	–0.89	.379	–0.13
	FU	(Both Positive vs. Neutral Control) × Time	–0.16 (0.13)	[–0.47, 0.15]	24.32	–1.24	.228	0.11
		(Both Positive vs. Both 50/50) × Time	–0.10 (0.15)	[–0.45, 0.26]	22.18	–0.66	.517	0.07
		(Both 50/50 vs. Neutral Control) × Time	–0.06 (0.13)	[–0.37, 0.25]	24.43	–0.48	.633	0.04
		(Positive + Negation vs. Positive) × Time	–0.17 (0.22)	[–0.70, 0.36]	21.63	–0.78	.446	–0.36
		(50/50 Blocked vs. 50/50 Random) × Time	0.27 (0.14)	[–0.06, 0.61]	31.68	1.94	.062	0.16
Growth mindset	TR	(Both Positive vs. Neutral Control) × Time	0.04 (0.03)	[–0.03, 0.12]	39.36	1.37	.180	0.20
		(Both Positive vs. Both 50/50) × Time	–0.01 (0.03)	[–0.08, 0.06]	41.55	–0.17	.863	–0.02
		(Both 50/50 vs. Neutral Control) × Time	0.05 (0.03)	[–0.03, 0.12]	36.71	1.47	.150	0.22
		(Positive + Negation vs. Positive) × Time	0.03 (0.05)	[–0.08, 0.13]	36.12	0.59	.560	0.12
		(50/50 Blocked vs. 50/50 Random) × Time	–0.04 (0.04)	[–0.14, 0.07]	39.14	–0.79	.432	–0.16
	FU	(Both Positive vs. Neutral Control) × Time	0.12 (0.15)	[–0.24, 0.47]	23.41	0.79	.439	0.33
		(Both Positive vs. Both 50/50) × Time	0.10 (0.17)	[–0.31, 0.51]	21.33	0.59	.559	0.09
		(Both 50/50 vs. Neutral Control) × Time	0.01 (0.12)	[–0.28, 0.31]	26.77	0.11	.911	0.23
		(Positive + Negation vs. Positive) × Time	–0.02 (0.17)	[–0.43, 0.39]	26.54	–0.12	.905	0.10
		(50/50 Blocked vs. 50/50 Random) × Time	0.31 (0.16)	[–0.07, 0.70]	28.72	1.95	.062	0.19
Optimism	TR	(Both Positive vs. Neutral Control) × Time	0.06 (0.03)	[–0.01, 0.13]	39.05	1.86	.070	0.24
		(Both Positive vs. Both 50/50) × Time	0.07 (0.03)	[0.01, 0.14]	41.79	2.50	.016*	0.31
		$\text{Time}_{\text{Both Positive}}$	0.17 (0.02)	[0.11, 0.22]	34.25	7.57	< .001**	0.72
		$\text{Time}_{\text{Both 50/50}}$	0.09 (0.02)	[0.04, 0.14]	39.89	4.29	< .001**	0.38
		(Both 50/50 vs. Neutral Control) × Time	–0.02 (0.02)	[–0.07, 0.04]	80.15	–0.70	.484	–0.07
		(Positive + Negation vs. Positive) × Time	0.05 (0.04)	[–0.05, 0.15]	38.45	1.13	.266	0.22
		(50/50 Blocked vs. 50/50 Random) × Time	0.03 (0.04)	[–0.05, 0.11]	75.97	0.81	.420	0.12

(continued)

Table 2. *(continued)*

Outcome	Phase	Effect	*b* (*SE*)	97.5% CI	*df*	*t*	*p*	*d*[a]
	FU	(Both Positive vs. Neutral Control) × Time	0.17 (0.15)	[−0.18, 0.53]	22.52	1.16	.258	0.43
		(Both Positive vs. Both 50/50) × Time	−0.22 (0.15)	[−0.58, 0.15]	22.19	−1.43	.166	0.09
		(Both 50/50 vs. Neutral Control) × Time	0.39 (0.10)	[0.16, 0.62]	34.14	4.03	< .001**	0.33
		$\text{Time}_{\text{Both 50/50}}$	0.29 (0.09)	[0.08, 0.49]	24.12	3.33	.003**	0.67
		$\text{Time}_{\text{Neutral Control}}$	−0.10 (0.09)	[−0.32, 0.12]	21.88	−1.13	.273	0.34
		(Positive + Negation vs. Positive) × Time	−0.16 (0.22)	[−0.68, 0.36]	21.97	−0.74	.469	0.04
		(50/50 Blocked vs. 50/50 Random) × Time	−0.03 (0.21)	[−0.52, 0.47]	22.65	−0.13	.901	0.09

Note: Separate models were fit for each outcome and reference group. Each model contained the fixed effects of condition, time_{TR} (TR phase trajectory), time_{FU} (FU phase trajectory), the Condition × Time_{TR} interaction, and the Condition × Time_{FU} interaction; a random intercept; and a random slope for time_{TR}. 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 data set). The combined-level data set, with condition coded in three levels (both positive, both 50/50, neutral control), was used to test interactions contrasting (a) both positive conditions with the neutral control condition, (b) both positive conditions with both 50/50 conditions, and (c) both 50/50 conditions with the neutral control condition. The separate-level data set, 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. CI = confidence interval; TR = training; FU = follow-up.

[a]For TR phase, *d* is the standardized mean difference between the given conditions at Session 4. For FU phase, *d* is the standardized mean difference at FU and is computed from both the Condition × Time_{TR} and Condition × Time_{FU} interaction effects.

*$p < .025$. **$p < .005$.

Growth mindset. Showing findings similar to those for depression and anxiety symptoms, ITT participants in all five conditions significantly improved in growth mindset during training (bs = 0.08–0.15, ps < .001, ds = 0.36–0.66; Table 1). No condition significantly changed from Session 4 to follow-up, and no significantly different changes emerged between the conditions we compared during training or from Session 4 to follow-up (Table 2; see Fig. S8 in the Supplemental Material).

Optimism. ITT participants in all five conditions significantly improved in optimism during training (bs = 0.09–0.19, ps < .001, ds = 0.37–0.79), and no individual condition significantly changed from Session 4 to follow-up (Table 1). As expected, during training, the two positive conditions improved significantly more than the two 50/50 conditions (b = 0.07, p = .016, d = 0.31), with no significant slope difference between the positive conditions and the 50/50 conditions from Session 4 to follow-up, yielding a between-groups effect size d from baseline to follow-up of 0.09 (Table 2; Fig. 2). Although no significant slope difference emerged between the 50/50 conditions and the neutral control condition during training, from Session 4 to follow-up the 50/50 conditions together continued to significantly improve (b = 0.29, p = .003) and significantly improved relative to the neutral control condition (b = 0.39, p < .001), yielding a between-groups effect size d from baseline to follow-up of 0.33. No significant slope difference emerged between the positive conditions and the neutral control condition during training or from Session 4 to follow-up.

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 showed iatrogenic effects based on the established criterion, pointing to the intervention's safety.

Discussion

The present randomized controlled trial evaluated a brief online CBM intervention to train less negative and more positive episodic prediction in a large sample of community participants with relatively negative expectancies, a transdiagnostic cognitive process common in emotional disorders. As hypothesized, during training, ITT participants in the positive conditions improved significantly more in positive expectancy bias, negative expectancy bias, and self-efficacy than participants in the neutral control condition; the improvements in bias were also significantly greater than those for participants in the 50/50 conditions. ITT participants in the positive conditions also improved significantly more in optimism than participants in the 50/50 conditions during training. ITT participants in all conditions generally improved on all outcomes during training and, despite some losses in training gains for positive expectancy bias in two conditions from Session 4 to follow-up, showed overall improvement from baseline to follow-up. Unexpectedly, participants in positive conditions did not improve in anxiety, depression, or growth mindset significantly more than control or 50/50 participants. In addition, no significantly different changes emerged between the neutral control and 50/50 conditions during training, although 50/50 participants improved significantly more in optimism than control participants during follow-up. Moreover, no significantly different changes emerged between the two 50/50 conditions or between the two positive conditions during training or follow-up in the ITT sample. PP participants (training completers) had 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 and/or 50/50 conditions) in the present study are broadly consistent with the results of Namaky et al. (2021). 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 relative to the 50/50 conditions, with no evidence of superiority for increasing growth mindset relative to the control or 50/50 conditions. Namaky et al. found the converse: superiority for increasing growth mindset relative to the control condition but not for increasing optimism relative to the control or 50/50 conditions. In addition to finding superiority for increasing positive bias, the present study found superiority for decreasing negative bias; improvements in the positive conditions surpassed not only those in the control condition but also those in the 50/50 conditions. The superior increases in positive expectancy bias in the positive conditions relative to the control (ITT: d = 0.80; PP: d = 0.58) and 50/50 (ITT: d = 0.63; PP: d = 0.44) conditions during training were medium to large in size. These effects are larger than the average pre-post (nonsignificant) increase in positive interpretation bias in benign conditions relative to

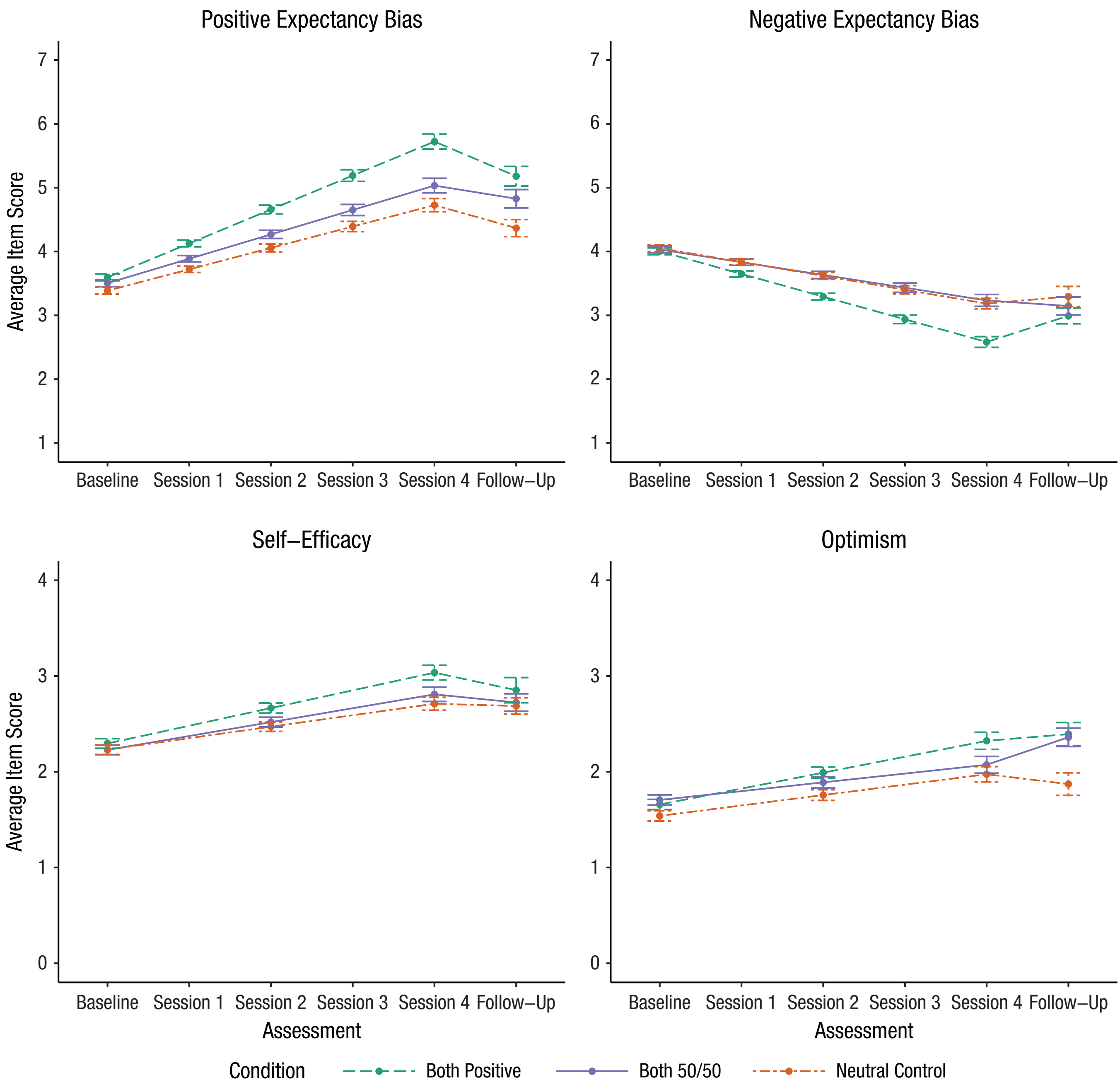

Fig. 2. Linear spline estimated means over time by condition for intent-to-treat sample. Means (±1 *SE*) estimated from the linear spline multilevel model based on the combined-level data set with neutral control as the reference group are shown. Estimates were computed from each imputed data set and then pooled following Rubin's rules using the *testConstraints* function of the *mitml* package (Version 0.4-1; Grund et al., 2021) in R (R Core Team, 2021). Plots were generated with the *ggplot2* (Version 3.3.3; Wickham et al., 2021) and *cowplot* (Version 1.1.1; Wilke, 2020) packages. Estimates for self-efficacy and optimism are not shown at Sessions 1 and 3 because these outcomes were not administered at these assessment points. Plots for anxiety and depression symptoms and growth mindset are shown in Figure S8 in the Supplemental Material available online.

neutral (including 50/50) conditions (d = 0.31, 95% CI [–0.19, 0.81]) found in a meta-analysis of different forms of CBM-I completed by anxious, depressed, and healthy samples (Menne-Lothmann et al., 2014). The present study and Namaky et al. suggest that resolving ambiguous future scenarios with mostly positive endings improves expectancy bias and self-efficacy, with mixed findings for growth mindset and optimism. Notably, we believe it is unlikely that demand characteristics can explain these condition effects given that the study rationale was described to all participants identically at consent, participants were not told their condition until debriefing, and we found no evidence that participants could tell whether they were in the active conditions (for details, see Section S1.16 in the Supplemental Material).

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 competing hypothesis that it harms by reinforcing negative associations (Ouimet et al., 2009). Likewise, 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 typically 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. (2021) 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. In addition, 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. (2021) 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.) Our results are consistent with a meta-analysis of different forms of CBM-I completed by anxious, depressed, and healthy samples (Menne-Lothmann et al., 2014), which found that the average difference in pre-post negative mood between benign conditions and neutral (including 50/50) conditions was nonsignificant ($d = 0.03$, 95% CI [−0.19, 0.25]), pointing to the need to strengthen these interventions for clinical applications. That said, a more recent meta-analysis of different forms of CBM-I found superior improvements in benign conditions relative to 50/50 (not including neutral) conditions for anxiety ($d = -0.31$, 95% CI [−0.53, −0.09]) and comorbid depression ($d = -0.47$, 95% CI [−0.79, −0.15]) in anxious ITT participants (Fodor et al., 2020).

Of course, because neither the present study nor Namaky et al. (2021) recruited on the basis of symptoms (in contrast to the studies included in the Fodor et al., 2020, meta-analysis, which included only trials with anxious or depressed participants), 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. (However, the analyses would lack random assignment to condition within each subgroup and have reduced power, and a meta-analysis of different forms of CBM-I completed by anxious, depressed, and healthy samples did not find that symptom presence significantly moderated the average effect of benign CBM-I conditions on pre-post negative mood [Menne-Lothmann et al., 2014].)

Another possible explanation for the lack of differential symptom improvement by condition concerns the analysis of anxiety and depression symptoms at the single-disorder, rather than the transdiagnostic, level (despite the intervention's transdiagnostic design). In an exploratory analysis, when Namaky et al. (2021) 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 (using separate models per study phase, before we later used linear splines) combined into one composite measure (the four-item sum of the PHQ-4) again found comparable improvement across conditions (see Sections S1.13 and S2.9 and Tables S11 and S12 in the Supplemental Material).

Feasibility of online CBM for episodic prediction

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 episodic prediction 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 intervention studies), transdiagnostic sample of 1,221 adults from 39 countries enrolled in the study for no payment; 79.5% across conditions started the first training session, 43.8% started the second session, 32.1% started the third session, 25.0% started the fourth session, and 12.3% started the 1-month follow-up assessment. Although we seek to reduce training dropout and loss to follow-up, the training dropout rate is lower than those of other free, web-based, self-help, at-home interventions that provide no payment or therapist contact for self-selected participants from the community. For example, only 10% of enrolled participants complete a second module at MoodGym, a cognitive behavioral website for depression (Batterham et al., 2008), and only 12% of enrolled participants complete the fourth session at the Panic Program, a cognitive behavioral website for panic (Farvolden et al., 2005). Dropout rates tend to be lower in web-based interventions that include even minimal therapist contact (Melville et al., 2010) or in randomized trials that tend to involve assessor contact or other factors (Christensen et al., 2009). 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, 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 training, who may need additional resources to maintain engagement.

Limitations

The present study's findings must be viewed in light of its limitations. First, the internal consistency of the expectancy bias measures was unacceptable, 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 et al., 2017), the bias results should be viewed with caution.

As others have discussed in the field of attentional bias (e.g., McNally, 2019), which has similarly encountered low internal consistency for measures of attentional bias to threat (e.g., dot-probe task), poor reliability can threaten inferences about *individuals*. For example, if eligible participants' relative expectancy bias index scores were more extreme than those of ineligible participants because of random measurement error, then it is possible that regression to the mean could explain some improvement in expectancy bias across conditions (see Rodebaugh et al., 2016). However, measures with lower reliability can still reveal average differences between *groups* (De Schryver et al., 2016; MacLeod et al., 2019), and repeated assessments can better reflect each participant's pattern of responses, which may also genuinely vary within person (MacLeod et al., 2019). Given that participants were randomly assigned to condition and that each condition's mean slope trajectory during training was estimated from expectancy bias scores at five assessment points, regression to the mean cannot account for differential improvement between conditions in expectancy bias during training, which occurred in hypothesized directions.

Still, because we sought to reduce symptoms in part through improving expectancy biases, if these biases did not actually improve despite improvements in the scores intended to measure them, which is possible if the scores are invalid, this may also explain the lack of differential symptom improvement between conditions (see MacLeod & Grafton, 2016). This situation highlights a 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. (2021) gives us greater confidence in our findings, ongoing construct validation is required, and we join others in calling for greater reporting of reliability and validity evidence, including for behavioral measures of cognitive processes (Parsons et al., 2019).

Second, participants were primarily female, White, and highly educated, which raises questions about the generalizability of the present results and highlights the potential need to culturally adapt behavioral intervention technologies to improve their reach for underserved samples (Ramos & Chavira, 2022).

Conclusion

Targeting transdiagnostic processes with technological interventions holds promise for improving well-being on a large scale. The present study is the first to target negative episodic prediction in community adults with a brief online CBM intervention and to show that doing so is feasible at scale, shifts expectancy bias, and improves outlook (although does not uniquely reduce symptoms). Future work is needed to develop more valid measures of bias, improve efficacy, and thereby advance the availability of evidence-based interventions for promoting psychological health.

Transparency

Action Editor: Colin MacLeod
Editor: Kenneth J. Sher
Author Contributions

B. A. Teachman developed the study concept. B. A. Teachman, L. E. Barnes, M. Boukhechba, D. Zhang, and D. H. Funk contributed to the study design and collected the data. J. W. Eberle, M. Boukhechba, and J. Sun performed the data analysis and interpretation. J. W. Eberle drafted the manuscript, and B. A. Teachman provided critical revisions. All of the authors approved the final manuscript for submission.

Declaration of Conflicting Interests

The authors declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

Grants awarded to B. A. Teachman from the National Institute of Mental Health (R34MH106770, R01MH113752) and the John Templeton Foundation (Science of Prospection Research Award) funded this work.

Open Practices

This article has received the badges for Open Data, Open Materials, and Preregistration. More information about the Open Practices badges can be found at http://www.psychologicalscience.org/publications/badges.

ORCID iDs

Jeremy W. Eberle https://orcid.org/0000-0003-0119-2879
Jianhui Sun https://orcid.org/0000-0003-0032-3646
Diheng Zhang https://orcid.org/0000-0001-8878-8952
Laura E. Barnes https://orcid.org/0000-0001-8224-5164
Bethany A. Teachman https://orcid.org/0000-0002-9031-9343

Acknowledgments

We thank Xin Tong; members of the Program for Anxiety, Cognition, and Treatment at the University of Virginia; and Judy Lightfoot for their comments on earlier drafts of this manuscript and members of the MindTrails Project research team for their support and feedback throughout the study. We also thank Craig Enders, Alan Feingold, Simon Grund, and Matteo Quartagno for their correspondence about statistical analyses.

Study data, analysis code, and materials are available at https://osf.io/jp5ws (Eberle et al., 2021). Following the preregistration (https://osf.io/jrst6), in the present study, we analyzed data collected from May 3, 2017, through September 9, 2018, for participants who enrolled on or before March 27, 2018. For the website code, see https://github.com/TeachmanLab/MindTrails/tree/templeton_release. This article is based on the master's thesis completed by Eberle (2019).

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/21677026221103128

Notes

1. 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.
2. Given the low rate of item-level missingness (see Missing Data Handling section), the disadvantages of listwise deletion (Enders, 2010, pp. 39–40) did not outweigh its convenience for assessing internal consistency.
3. In R (R Core Team, 2020), we computed standardized Cronbach's alpha (see Table S4 in the Supplemental Material available online) using the *psych* package (Version 1.9.12.31; Revelle, 2019) before learning that McDonald's omega total is recommended instead (Dunn et al., 2014), which we computed using the *MBESS* package (Ver. 4.7.0; Kelley, 2020; for details, see Section S1.5 in the Supplemental Material).
4. In an earlier version of this article, we did not include condition as a Level 2 predictor in the imputation model. We also included only one linear trajectory from baseline to follow-up in the imputation model and then used separate analysis models for each study phase—training and follow-up. Because more data were observed during training than at follow-up, the imputed data at follow-up were overly influenced by the training trajectory, overestimating improvement during follow-up. Although nearly all training effects were robust to our new linear spline models, we present full results of these earlier analyses in Tables S6 through S12 and Figures S1 through S7 in the Supplemental Material. For more details about this original approach to our statistical analyses, see Section S1.13 in the Supplemental Material.

References

Altman, D. G. (2009). Missing outcomes in randomized trials: Addressing the dilemma. *Open Medicine*, *3*(2), e51–e53.

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

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.

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), 179–183. https://doi.org/10.1037/0021-843X.96.3.179

Becker, K. D., Chorpita, B. F., & Daleiden, E. L. (2011). Improvement in symptoms versus functioning: How do our best treatments measure up? *Administration and Policy in Mental Health and Mental Health Services Research*, *38*, 440–458. https://doi.org/10.1007/s10488-010-0332-x

Blackwell, S. E., Browning, M., Mathews, A., Pictet, A., Welch, J., Davies, J., Watson, P., Geddes, J. R., & 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. https://doi.org/10.1177/2167702614560746

Blackwell, S. E., & Holmes, E. A. (2010). Modifying interpretation and imagination in clinical depression: A single case series using cognitive bias modification. *Applied Cognitive Psychology*, *24*, 338–350. https://doi.org/10.1002/acp.1680

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. https://doi.org/10.1016/j.brat.2017.10.010

Borkovec, T. D., & Nau, S. D. (1972). Credibility of analogue therapy rationales. *Journal of Behavior Therapy and Experimental Psychiatry*, *3*(4), 257–260. https://doi.org/10.1016/0005-7916(72)90045-6

Bowler, J. O., Mackintosh, B., Dunn, B. D., Mathews, A., Dalgleish, T., & Hoppitt, L. (2012). A comparison of cognitive bias modification for interpretation and computerized cognitive behavior therapy: Effects on anxiety, depression, attentional control, and interpretive bias. *Journal of Consulting and Clinical Psychology, 80*(6), 1021–1033. https://doi.org/10.1037/a0029932

Carver, C. S., & Scheier, M. F. (2018). Generalized optimism. In G. Oettingen, A. T. Sevincer, & P. M. Gollwitzer (Eds.), *The psychology of thinking about the future* (pp. 214–230). The Guilford Press.

Center for Evidence-Based Practices. (2010). *Readiness ruler*. Case Western Reserve University. https://www.centerforebp.case.edu/resources/tools/readiness-ruler

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a New General Self-Efficacy Scale. *Organizational Research Methods*, *4*(1), 62–83. https://doi.org/10.1177/109442810141004

Christensen, H., Griffiths, K. M., & Farrer, L. (2009). Adherence in Internet interventions for anxiety and depression: Systematic review. *Journal of Medical Internet Research*, *11*(2), Article e13. https://doi.org/10.2196/jmir.1194

Clarke, P. J. F., Nanthakumar, S., Notebaert, L., Holmes, E. A., Blackwell, S. E., & MacLeod, C. (2014). Simply imagining sunshine, lollipops and rainbows will not budge the bias: The role of ambiguity in interpretation bias modification. *Cognitive Therapy and Research*, *38*, 120–131. https://doi.org/10.1007/s10608-013-9564-x

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. https://doi.org/10.1037/1082-989X.6.4.330

De Schryver, M., Hughes, S., Rosseel, Y., & De Houwer, J. (2016). Unreliable yet still replicable: A comment on LeBel and Paunonen (2011). *Frontiers in Psychology*, *6*, Article 2039. https://doi.org/10.3389/fpsyg.2015.02039

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. https://doi.org/10.1111/bjop.12046

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

Dweck, C. S., & Yeager, D. S. (2018). Mindsets change the imagined and actual future. In G. Oettingen, A. T. Sevincer, & P. M. Gollwitzer (Eds.), *The psychology of thinking about the future* (pp. 362–376). The Guilford Press.

Eberle, J. W. (2019). *Shifting negative prospection with online cognitive bias modification: A randomized controlled trial* [Master's thesis, University of Virginia]. University of Virginia LibraETD digital archive. https://doi.org/10.18130/v3-s1vt-9104

Eberle, J. W., Boukhechba, M., Sun, J., Zhang, D., Funk, D., Barnes, L., & Teachman, B. (2021, August 20). *Shifting episodic prediction with online cognitive bias modification: A randomized controlled trial* [Data, analysis code, materials]. https://osf.io/jp5ws

Enders, C. K. (2010). *Applied missing data analysis*. The Guilford Press.

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

Farvolden, P., Denisoff, E., Selby, P., Bagby, M., & Rudy, L. (2005). Usage and longitudinal effectiveness of a web-based self-help cognitive behavioral therapy program for panic disorder. *Journal of Medical Internet Research*, *7*(1), Article e7. http://doi.org/10.2196/jmir.7.1.e7

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. https://doi.org/10.1037/a0014699

Feingold, A. (2018). Time-varying effect sizes for quadratic growth models in multilevel and latent growth modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(3), 418–429. https://doi.org/10.1080/10705511.2018.1547110

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. https://doi.org/10.1177/1948550617693063

Fodor, L. A., Georgescu, R., Cuijpers, P., Szamoskozi, S., David, D., Furukawa, T. A., & Cristea, I. A. (2020). The effectiveness of cognitive bias modification interventions in anxiety and depressive disorders: A network meta-analysis. *The Lancet Psychiatry*, *7*(6), P506–P514. https://doi.org/10.1016/S2215-0366(20)30130-9

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

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. https://doi.org/10.1177/1094428117703686

Grund, S., Robitzsch, A., & Luedtke, O. (2021). *mitml: Tools for multiple imputation in multilevel modeling* [Computer software]. https://cran.r-project.org/package=mitml

Hohensee, N., Meyer, M. J., & Teachman, B. A. (2020). The effect of confidence on dropout rate and outcomes in online cognitive bias modification. *Journal of Technology in Behavioral Science*, *5*, 226–234. https://doi.org/10.1007/s41347-020-00129-8

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

Ji, J. L., Baee, S., Zhang, D., Calicho-Mamani, C. P., Meyer, M. J., Funk, D., Portnow, S., Barnes, L., & Teachman, B. A. (2021). Multi-session online interpretation bias training for anxiety in a community sample. *Behaviour Research and Therapy*, *142*, Article 103864. https://doi.org/10.1016/j.brat.2021.103864

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

Kazdin, A. E. (2001). Almost clinically significant (p < .10): Current measures may only approach clinical significance. *Clinical Psychology: Science and Practice*, *8*(4), 455–462. https://doi.org/10.1093/clipsy.8.4.455

Kelley, K. (2020). *MBESS: The MBESS R package* [Computer software]. https://cran.r-project.org/package=MBESS

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. https://doi.org/10.1097/01.MLR.0000093487.78664.3C

Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Löwe, B. (2009). An ultra-brief screening scale for anxiety and depression: The PHQ-4. *Psychosomatics*, *50*(6), 613–621. https://doi.org/10.1176/appi.psy.50.6.613

Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Löwe, B. (2010). The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: A systematic review. *General Hospital Psychiatry*, *32*(4), 345–359. https://doi.org/10.1016/j.genhosppsych.2010.03.006

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. https://doi.org/10.7326/0003-4819-146-5-200703060-00004

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. https://doi.org/10.1002/per.855

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

Löwe, B., Kroenke, K., & Grafe, K. (2005). Detecting and monitoring depression with a two-item questionnaire (PHQ-2). *Journal of Psychosomatic Research*, *58*, 163–171. https://doi.org/10.1016/j.jpsychores.2004.09.006

MacLeod, C., & Grafton, B. (2016). Anxiety-linked attentional bias and its modification: Illustrating the importance of distinguishing processes and procedures in experimental psychopathology research. *Behaviour Research and Therapy*, *86*, 68–86. https://doi.org/10.1016/j.brat.2016.07.005

MacLeod, C., Grafton, B., & Notebaert, L. (2019). Anxiety-linked attentional bias: Is it reliable? *Annual Review of Clinical Psychology*, *15*, 529–554. https://doi.org/10.1146/annurev-clinpsy-050718-095505

Maddux, J. E., & Kleiman, E. M. (2018). Self-efficacy. In G. Oettingen, A. T. Sevincer, & P. M. Gollwitzer (Eds.), *The psychology of thinking about the future* (pp. 174–198). The Guilford Press.

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

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

McGuire, J. T., Nassar, M. R., Gold, J. I., & Kable, J. W. (2014). Functionally dissociable influences on learning rate in a dynamic environment. *Neuron*, *84*(4), 870–881. https://doi.org/10.1016/j.neuron.2014.10.013

McNally, R. J. (2019). Attentional bias for threat: Crisis or opportunity? *Clinical Psychology Review*, *69*, 4–13. https://doi.org/10.1016/j.cpr.2018.05.005

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. https://doi.org/10.1016/j.jbtep.2011.02.012

Melville, K. M., Casey, L. M., & Kavanagh, D. J. (2010). Dropout from Internet-based treatment for psychological disorders. *British Journal of Clinical Psychology*, *49*, 455–471. https://doi.org/10.1348/014466509X472138

Menne-Lothmann, C., Viechtbauer, W., Höhn, P., Kasanova, Z., Haller, S. P., Drukker, M., van Os, J., Wichers, M., & Lau, J. Y. F. (2014). How to boost positive interpretations? A meta-analysis of the effectiveness of cognitive bias modification for interpretation. *PLOS ONE*, *9*(6), Article e100925. https://doi.org/10.1371/journal.pone.0100925

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

Murphy, S. E., O'Donoghue, M. C., Blackwell, S. E., Nobre, A. C., Browning, M., & Holmes, E. A. (2017). Increased rostral anterior cingulate activity following positive mental imagery training in healthy older adults. *Social Cognitive and Affective Neuroscience*, *12*, 1950–1958. https://doi.org/10.1093/scan/nsx120

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. https://doi.org/10.1016/j.psychres.2015.07.059

Namaky, N., Glenn, J. J., Eberle, J. W., & Teachman, B. A. (2021). Adapting cognitive bias modification to train healthy prospection. *Behaviour Research and Therapy*, *144*, Article 103923. https://doi.org/10.1016/j.brat.2021.103923

Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., Estabrook, R., Bates, T. C., Maes, H. H., & Boker, S. M. (2016). OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, *81*(2), 535–549. https://doi.org/10.1007/s11336-014-9435-8

Ouimet, A. J., Gawronski, B., & Dozois, D. J. (2009). Cognitive vulnerability to anxiety: A review and an integrative model. *Clinical Psychology Review*, *29*(6), 459–470. https://doi.org/10.1016/j.cpr.2009.05.004

Parsons, S., Kruijt, A., & Fox, E. (2019). Psychological science needs a standard practice of reporting the reliability of cognitive-behavioral measurements. *Advances in Methods and Practices in Psychological Science*, *2*(4), 378–395. https://doi.org/10.11772F2515245919879695

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. https://doi.org/10.1177/0146167299025002010

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2021). *nlme: Linear and nonlinear mixed effects models* [Computer software]. https://CRAN.R-project.org/package=nlme

Quartagno, M., & Carpenter, J. (2020). *jomo: A package for multilevel joint modelling multiple imputation* [Computer software]. 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. https://doi.org/10.1080/17439760902992365

Ramos, G., & Chavira, D. A. (2022). Use of technology to provide mental health care for racial and ethnic minorities: Evidence, promise, and challenges. *Cognitive and Behavioral Practice*, *29*(1), 15–40. https://doi.org/10.1016/j.cbpra.2019.10.004

R Core Team. (2020). *R: A language and environment for statistical computing* (Version 4.0.2) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

R Core Team. (2021). *R: A language and environment for statistical computing* (Version 4.1.0) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

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. https://doi.org/10.1007/s10608-016-9776-y

Revelle, W. (2019). *psych: Procedures for psychological, psychometric, and personality research* [Computer software]. https://cran.r-project.org/package=psych

Rodebaugh, T. L., Scullin, R. B., Langer, J. K., Dixon, D. J., Huppert, J. D., Bernstein, A., & Zvielli, A. (2016). Unreliability as a threat to understanding psychopathology: The cautionary tale of attentional bias. *Journal of Abnormal Psychology*, *125*(6), 850–851. https://doi.org/10.1037/abn0000184

Roepke, A. M., & Seligman, M. E. P. (2016). Depression and prospection. *British Journal of Clinical Psychology*, *55*(1), 23–48. https://doi.org/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. https://doi.org/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. https://doi.org/10.1177/1745691612474317

Staples, L. G., Dear, B. F., Gandy, M., Fogliati, V., Fogliati, R., Karin, E., Nielssen, O., & Titov, N. (2019). Psychometric properties and clinical utility of brief measures of depression, anxiety, and general distress: The PHQ-2, GAD-2, and K-6. *General Hospital Psychiatry*, *56*, 13–18. https://doi.org/10.1016/j.genhosppsych.2018.11.003

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. https://doi.org/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, USA*, *111*(52), 18414–18421. https://doi.org/10.1073/pnas.1417144111

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. https://doi.org/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. https://doi.org/10.1177/1745691613512659

Wickham, H., Chang, W., Henry, L., Pederson, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., & Dunnington, D., & RStudio. (2021). *ggplot2: Create elegant data visualisations using the grammar of graphics* [Computer software]. https://cran.r-project.org/package=ggplot2

Wilke, C. O. (2020). *cowplot: Streamlined plot theme and plot annotations for 'ggplot2'* [Computer software]. https://cran.r-project.org/package=cowplot