the british psychological society
promoting excellence in psychology

British Journal of Clinical Psychology (2021)

www.wileyonlinelibrary.com

Invited Article

Development of a cognitive bias modification intervention for anxiety disorders in primary care

Risa B. Weisberg[1,2,3], Meghan A. Gonsalves[4], Ramya Ramadurai[5], Howard Braham[6], Cara Fuchs[2,7] and Courtney Beard[5,8]*

[1]VA Boston Healthcare System, Massachusetts, USA
[2]Department of Psychiatry, Boston University School of Medicine, Massachusetts, USA
[3]Department of Family Medicine, Alpert Medical School, Brown University, Providence, Rhode Island, USA
[4]Neuroscience Graduate Program, Brown University, Providence, Rhode Island, USA
[5]Department of Psychiatry, McLean Hospital, Belmont, Massachusetts, USA
[6]Kabram, Inc, Burbank, California, USA
[7]Boston Medical Center, Massachusetts, USA
[8]Department of Psychiatry, Harvard Medical School, Boston, Massachusetts, USA

Objectives. There is a great need for low-intensity, scalable treatments in primary care, where most anxious patients first present for treatment. We describe Stage IA treatment development and a Stage IB feasibility trial of cognitive bias modification (CBM) for transdiagnostic anxiety in primary care.

Methods. The online intervention, Mental Habits, comprised eight sessions of a personalized CBM targeting attention and interpretation biases. Coaches assisted patients in using the website, monitored progress via a dashboard, and shared information with primary care providers. We evaluated Mental Habits in an open trial (N = 14) and a randomized controlled trial (RCT) (N = 40) in primary care patients with anxiety disorders.

Results. We compared results to a priori benchmarks of clinically meaningful outcomes. In the open trial, Mental Habits met feasibility, acceptability, and efficacy benchmarks. In the pilot RCT, there was greater dropout at one study site which ultimately closed. In the intent-to-treat analyses, Mental Habits met the benchmark for self-report, but not the interview measure of anxiety. Symptom Tracking did not meet the benchmark for self-report or interview measures of anxiety. In per-protocol analyses, Mental Habits exceeded the benchmark for both self-report and interview measures, whereas Symptom Tracking met the benchmark for self-report. Interpretation bias improved in the Mental Habits group, but not in Symptom Tracking. No effects were observed for attention bias.

Conclusion. The online CBM intervention demonstrated good acceptability and, when delivered at a stable primary care clinic, preliminary effectiveness in primary care. A larger RCT is warranted to test effectiveness.

DOI:10.1111/bjc.12281

Practitioner points

- A personalized, transdiagnostic Cognitive Bias Modification (CBM) intervention for anxiety in primary care is acceptable to primary care patients with social anxiety disorder, generalized anxiety disorder, and/or panic disorder /agoraphobia.
- With training and supervision from licensed mental health clinicians, bachelor's-level coaches can assist primary care patients to self-administer CBM.
- Offering a low-intensity, self-directed anxiety intervention in primary care can greatly expand the reach of anxiety treatment, with minimal need for additional resources.
- Interpretation bias may be an important clinical target for primary care patients with anxiety.

Most people with anxiety disorders seek treatment in primary care rather than specialty settings (Verhaak et al., 2009). Yet, most primary care patients with anxiety disorders do not receive appropriate treatment (Stein et al., 2004; Weisberg, Beard, Moitra, Dyck, & Keller, 2014). Thus, developing acceptable and widely scalable treatments for anxiety in primary care is prudent.

Pharmacotherapy is efficacious for anxiety disorders (e.g., Kimmel, Roy-Byrne, & Cowley, 2014), but many patients are reluctant to take medications and prefer non-pharmacological intervention (McHugh, Whitton, Peckham, Welge, & Otto, 2013; Mohlman, 2012). Cognitive-behavioural therapy (CBT) is efficacious for anxiety disorders (Chavira et al., 2014; van Dis et al., 2020; Zhang et al., 2019), but implementation is limited by cost, availability of trained providers (Kazdin & Blase, 2011), and stigma, particularly in underserved and marginalized communities (e.g., low-income and people of colour; Ofonedu, Belcher, Budhathoki, & Gross, 2017; Thompson, Bazile, & Akbar, 2004).

Computerized and smartphone delivered interventions may help bridge the gap. For example, the Intellicare suite of CBT smartphone apps reduced anxiety in primary care patients compared to a waitlist (Graham et al., 2020). This study used bachelor's-level coaches to monitor patient progress and enhance engagement. Such low-intensity interventions may reach more people and save intensive resources for those who need them. Primary care is well suited to low-intensity interventions, as primary care providers (PCPs) can monitor patient progress and increase treatment intensity as needed.

Another low-intensity anxiety intervention is Cognitive Bias Modification (CBM). CBM targets cognitive biases underlying anxiety disorders via repetitive training tasks. To date, most work in this area has focused on attention and interpretation biases due to their established role in emotional disorders (e.g., Goodwin, Yiend, & Hirsch, 2017; Hirsch et al., 2016). For example, Attention Bias Modification (ABM) targets attention bias – the tendency to selectively attend to threat stimuli, even when those stimuli are irrelevant to current goals. In an ABM dot probe task, patients repeatedly identify a probe (e.g., the letter 'E' or 'F') that replaces a neutral stimulus (e.g., neutral face) competing for attention with a threatening stimulus (e.g., angry face). ABM has shown promise as an intervention for a range of anxiety disorders, most notably general anxiety (e.g., Amir et al., 2009) and social anxiety (e.g., Amir et al., 2008). Meta-analyses of ABM initially showed reliable effects on anxiety, with small effect sizes (Beard, Sawyer, & Hofmann, 2012; Linetzky, Pergamin-Hight, Pine, & Bar-Haim, 2015; Mogoaşe, David, & Koster, 2014). However, more recent reviews call into question the effectiveness of ABM (Cristea, Kok, & Cuijpers, 2015; Fodor et al., 2020).

*Correspondence should be addressed to Courtney Beard, Department of Psychiatry, McLean Hospital, 115 Mill St., Belmont, MA 02478, USA (email: cbeard@mclean.harvard.edu).

Cognitive Bias Modification for Interpretation (CBM-I) attempts to modify interpretation bias, the tendency to interpret ambiguous cues in a negative or threatening manner, which has also been implicated in anxiety disorders (see Hirsch et al., 2016). For example, the CBM-I version of the Word Sentence Association Paradigm presents ambiguous situations and words representing benign or threatening interpretations of those situations (see Beard & Amir, 2008). Patients make judgements about the relatedness of the word and sentence and are provided corrective feedback reinforcing a benign interpretive style. CBM-I has been effective in shifting interpretation and symptoms in general anxiety (e.g., Hirsch et al., 2018) and social anxiety (Amir & Taylor, 2012). To date, fewer studies have applied this intervention to panic disorder (e.g., Beard et al., 2016), although we would expect CBM-I to be a relevant treatment because it targets the catastrophic misinterpretation of physical sensations. In contrast to ABM, CBM-I has consistently shown promise as an intervention for anxiety with larger effects than ABM (Fodor et al., 2020; Hallion & Ruscio, 2011; Liu, Li, Han, & Liu, 2017).

CBM has potential for broad scalability due to its computerized or app-delivery, without requiring a trained therapist or the patient to complete challenging assignments on their own. Thus, CBM may be well suited as a low-intensity intervention delivered in primary care. A prior qualitative study suggested that CBM may be an acceptable approach to primary care patients with social anxiety (Beard, Weisberg, & Primack, 2012). A different type of CBM, concreteness training, reduced rumination in depressed primary care patients compared to treatment as usual and relaxation training (Watkins et al., 2012), but it was not superior to relaxation in reducing depression symptoms. To our knowledge, no other studies have tested CBM in primary care.

We sought to develop and pilot test a CBM intervention for primary care patients with a range of anxiety disorders. We tested a combined ABM/CBM-I intervention to maximize impact on cognitive biases and ultimately on anxiety. A handful of prior studies have tested combined ABM and CBM-I interventions (e.g., Beard, Weisberg, & Amir, 2011; Naim, Kivity, Bar-Haim, & Huppert, 2018; Yeung & Sharpe, 2019) in hopes that a combined intervention may lead to larger clinical benefits both due to additive effects, as well as synergistic effects of targeting the relatively 'bottom-up' attentional processes and 'top-down' interpretation processes.

Consistent with the Stage Model of Behavioral Therapies Research (Onken, Carroll, Shoham, Cuthbert, & Riddle, 2014; Rounsaville, Carroll, & Onken, 2001), we developed the intervention via Stage 1A intervention generation and refinement. We created a CBM intervention appropriate for primary care patients with social anxiety disorder (SAD), generalized anxiety disorder (GAD), and/or panic disorder/agoraphobia (PD/A) which we pilot tested in an open trial (Study 1). In Study 2, we conducted a Stage IB feasibility and pilot randomized control trial (RCT), in which we compared the CBM intervention to a symptom monitoring condition. Our aims were to 1) develop a personalized, transdiagnostic, CBM intervention and the methods for linkage to primary care, and 2) compare feasibility, acceptability, and efficacy outcomes to a priori benchmarks. We hypothesized that the CBM and Symptom Tracking arms would meet our a priori benchmarks for feasibility and acceptability. We also hypothesized that CBM would meet benchmarks for clinical outcomes, whereas Symptom Tracking would not.

STUDY 1

In Stage 1A, we developed the CBM intervention, 'Mental Habits', by incorporating stakeholder feedback and adapting existing CBM protocols to the primary care population and setting. To ensure that Mental Habits would be appropriate for patients with a range of comorbid anxiety disorders, we developed a means to personalize stimuli. Next, we developed methods for linking Mental Habits to primary care, including a dashboard for monitoring patient progress, communication of patient progress with PCPs through the Electronic Health Record (EHR), identification of patients requiring further care, and creation of a guide for coaches. Finally, we conducted an open trial to refine the intervention and collect preliminary data.

Study 1 Methods

Mental habits

Mental Habits is an online intervention including: 1) two brief videos describing the intervention and rationale; 2) a personalization process to select stimuli; 3) eight, 20-minute sessions of ABM followed by CBM-I; 4) monitoring of anxiety and depression symptoms, and 5) weekly 10-minute check-ins with a coach.

Participants were asked to complete two CBM sessions per week, separated by at least one day, consistent with prior protocols (e.g., Amir et al., 2009; Beard et al., 2016). Participants completed the first session in primary care with the coach. After this session, participants could complete remaining sessions in primary care or at home.

Personalization

To personalize the CBM stimuli, patients completed a checklist of life circumstances (e.g., employment, marital status, etc.) and self-report measures of symptom severity. To assess GAD symptoms, we administered the Penn State Worry Questionnaire (PSWQ; Meyer, Miller, Metzger, & Borkovec, 1990), a 16-item questionnaire that measures traits of worry related GAD on a Likert-type scale. To assess social anxiety, we administered the Fear subscale of the Liebowitz Social Anxiety Scale (LSAS-F; Liebowitz, 1987), a measure of the frequency of anxiety in 24 social situations. To assess panic symptoms, we administered the Panic Disorder Severity Scale (PDSS; Shear et al., 1997), a 7-item scale assessing the frequency, distress, and impairment caused by PD/A symptoms. An algorithm used these data to select personalized stimuli for CBM (e.g., for participants endorsing predominantly SAD symptoms, the program selected faces displaying disgust for ABM and sentences related to social judgement for CBM-I (see Table 1)). The program also selected CBM-I stimuli relevant to each participant's life. For example, a patient who was employed might view work-related situations, one with children might view parenting situations, etc.

Attention bias modification

At each session, participants completed 128 trials of a dot probe task. We selected images of people from various racial and ethnic backgrounds making fear, anger, and disgust facial expressions from the NimStim Face Stimuli Set (Tottenham et al., 2009). Patients were asked to focus on a fixation cross, which appeared at the centre of the screen for 500ms. Next, two faces (one threat, one neutral) appeared for 500ms simultaneously, one centred on the top half of the screen, the other centred on the bottom half. Each face

Table 1. GAD, SAD, and PD CBM attention and interpretation stimuli

Anxiety disorder	Mental habits attention stimuli	Example mental habits interpretation stimuli	
Generalized anxiety disorder	Angry faces	Ambiguous sentence	You are asked to assume a new responsibility at work
		Threat word	Inadequate
		Benign word	Deserving
Social anxiety disorder	Disgust faces	Ambiguous sentence	People stare at you at a restaurant
		Threat word	Unattractive
		Benign word	Attractive
Panic disorder	Fear faces	Ambiguous sentence	While exercising, you feel your heart rate increase.
		Threat word	Problem
		Benign word	Optimal

Note. This table provides examples of Mental Habits Attention and Interpretation stimuli used in both the Open Trial and Randomized Control Trial for each anxiety diagnosis. Attention stimuli were anger, fear, or disgust faces from the NimStim dataset (Tottenham et al., 2009), and interpretation stimuli were taken from the Word Sentence Association Paradigm (WSAP; Beard & Amir, 2009). Note: Individuals endorsing multiple types of anxiety saw a mixture of stimuli types.

appeared randomly in either the top or bottom position. Afterwards, a probe replaced the neutral face. Participants indicated if the probe was an 'E' or an 'F' as quickly and accurately as possible by pressing the left arrow key for an 'E' and the right arrow key for an 'F'. The arrow keys were temporarily labelled with 'E' and 'F' stickers. This paradigm is designed to train participants to shift their attention towards the neutral face and away from the threatening face.

Interpretation bias modification

In each session, participants completed 130 trials of a personalized CBM-I version of the Word-Sentence Association Paradigm (WSAP; Beard & Amir, 2009; Gonsalves, Whittles, Weisberg, & Beard, 2019). Participants were asked to focus on a fixation cross in the middle of the screen that appeared for 500ms. After the cross disappeared, a word appeared in its place for 500ms. The word was either a threat (e.g., '*cancer*') or benign (e.g., '*height*') interpretation of a sentence that followed (e.g., '*The doctor measures your growth'*.). The sentence appeared on the screen until the participant pressed a computer key to indicate whether they thought the word was related to the sentence. Participants pressed the left arrow key to indicate 'Yes', the word and sentence were related, and the right arrow key to indicate 'No', the word and sentence were not related. The arrow keys were labelled with 'Y' for 'Yes' and 'N' for 'No'. 'You are correct!' appeared if they endorsed the benign interpretation and rejected the threatening one, and 'Let's try another one' appeared if they endorsed the threatening interpretation and rejected the benign one. This task is designed to train participants to accept benign interpretations off ambiguous situations.

Coaches

Undergraduate research assistants (trained and supervised by licensed psychologists) provided support with the online program, serving as a link between the participant,

CBM, and the patient's PCP. In the first Mental Habits session, the coach oriented the participant to the intervention. Thereafter, the coach conducted 10-minute weekly check-ins with each participant. During check-ins, participants completed the GAD-7 (Spitzer, Kroenke, Williams, & Löwe, 2006) and PHQ-9 (Kroenke, Spitzer, & Williams, 2001) online to monitor symptoms. Graphs of accuracy and reaction time on the CBM tasks, and GAD-7 and PHQ-9 scores, over time were created by the program and discussed during check-ins. Coaches used these graphs and motivation enhancement techniques to encourage continued treatment engagement. The coach also inquired about recent medication changes, medical visits, and major life events. Coaches were necessarily not blind to treatment arm and were aware of the goals of the study.

Coaches monitored patient safety via the dashboard: Suicidal ideation (score > 0 on PHQ-9 item 9) and clinical deterioration (increase of 5 or more points on the GAD-7 or PHQ-9) triggered an alert, and a licensed clinician conducted a risk assessment. Finally, coaches communicated participants' progress to PCPs through the EHR.

Participants and setting

Participants were recruited from a university-affiliated, hospital-based, family medicine practice in urban/suburban Rhode Island. The clinic served a racially, ethnically, and economically diverse population. The practice had a small, integrated behavioural health service staffed by advanced mental health trainees.

Study recruitment methods included the following: 1) referrals from providers in the practice; 2) waiting room recruitment; and 3) brochures in the clinic. Interested patients

Table 2. Demographic and diagnostic characteristics

	Mental Habits (Open Trial) $N = 14$		Mental Habits (RCT) $N = 19$		Symptom Tracking (RCT) $N = 21$	
Characteristic	*n*	%	*n*	%	*n*	%
Age in years (*M, SD*)	49.79 (17.37)		42.56 (14.16)[†]		41.10 (15.88)[‡]	
Male	2	14.29%	5	26.32%	4	19.05%
Education (≥ 12 years)	12	85.71%	14	73.68%	19	90.48%
Hispanic/Latinx	1	7.14%	0	0%	5	23.81%
Race						
Asian/Pacific	0	0%	0	0%	0	0%
Black	0	0%	3	15.79%	1	4.76%
Multiracial	0	0%	2	10.53%	4	19.05%
Native American	0	0%	0	0%	0	0%
White	12	85.71%	14	73.68%	13	61.91%
Not reported	2	14.29%	0	0%	3	14.29%
Anxiety diagnoses						
GAD	10	71.43%	12	63.16%	15	71.43%
SAD	5	35.71%	10	52.63%	11	52.38%
PD/A	10	71.43%	11	57.89%	8	38.10%

Note. GAD = Generalized Anxiety Disorder; SAD = Social Anxiety Disorder; PD/A = Panic Disorder with or without Agoraphobia. Diagnostic percentages add up to greater than 100%, due to comorbidity
[†]One participant in the Mental Habits condition did not provide their age.; [‡]One participant in the Symptom Tracking condition did not provide their age.

were screened for initial eligibility in-person or by phone, depending on referral method. Patients were eligible for a pre-treatment assessment if they were a patient of the practice, had a GAD-7 score ≥ 10, were at least 18 years old, and were fluent in English. Inclusion criteria assessed at the pre-treatment assessment included a primary DSM-IV diagnosis of SAD, GAD, and/or PD/A established by a PhD-level psychologist using the Structured Clinical Interview for the DSM-IV (First et al., 2002). Exclusion criteria included current suicidal ideation, psychotic symptoms, moderate-to-severe cognitive impairment as identified by the PCP, current weekly or biweekly psychotherapy in the community (patients who met with a primary care-behavioural health provider at the primary care practice were not excluded, as these meetings tended be very brief and to occur less regularly than biweekly), and change in pharmacological treatments during the eight weeks prior to the first treatment session.

Of the 55 patients who consented to be screened, 31 (56.36%) were ineligible due to medication instability, current psychotherapy, or GAD-7 score < 10. Of the 24 patients invited for a pre-treatment assessment, nine were lost to follow-up or no longer interested, and one was excluded after assessment due to inability to read. Ultimately, 14 participants enrolled. They were predominantly female and White, with a mean age of 50 (see Table 2).

Measures

Primary outcomes were feasibility and acceptability. The Credibility Expectancy Questionnaire (CEQ; Devilly & Borkovec, 2000) was administered immediately after the first session to assess treatment credibility. The Client Satisfaction Questionnaire (CSQ-8; Larsen, Attkinson, Hargreaves, & Nguyen, 1979) was administered at post-treatment, as a measure of acceptability.

To pilot methods for the subsequent Stage IB pilot RCT, we administered symptom measures. The *primary symptom measure* was the Hamilton Anxiety Rating Scale (HARS; Hamilton, 1959), a 14-item clinician-administered measure of anxiety severity. A master's-level social worker or doctoral level psychologist, who received HARS training from a study investigator (RW), administered the HARS at pre-, mid-, and post-treatment assessments. These assessors were blind to treatment condition and were not involved in the study outside of conducting assessments. Inter-rater reliability in prior work has been reported as an intra-class correlation coefficient of 0.74–0.96 (Bruss, Gruenberg, Goldstein, & Barber, 1994). *Anxiety symptoms* were also assessed via participant report with the GAD-7 (Spitzer et al., 2006), the most widely used and validated anxiety measure in primary care (Mulvaney-Day et al., 2018). GAD-7 scores range from zero to 21 with scores of five, 10, and 15 representing mild, moderate, and severe anxiety symptoms, respectively. The GAD-7 was used to establish initial eligibility for a pre-treatment assessment (scores > 10) and completed at pre-treatment, sessions two, four, six, and eight, and the post-treatment assessment to assess changes in anxiety severity over time.

Procedure

Participants were asked to complete Mental Habits and to attend follow-up assessments at the primary care site for mid-treatment and post-treatment time points. They were compensated $30 for the pre-treatment, $20 for the mid-treatment, and $30 for the post-treatment assessment, which was scheduled for one week after intervention completion.

Study 1 Results

One participant was lost to follow-up after the pre-treatment assessment and completed no Mental Habits sessions nor the post-treatment assessment. No participants experienced clinical deterioration or safety concerns warranting referral to additional treatment during the study.

As this was a small, uncontrolled, open trial, we established a priori benchmarks (Table 3) for determining intervention feasibility, acceptability, and preliminary effectiveness. These benchmarks were selected based on prior CBM trials in real-world clinical settings (e.g., Beard, Ramadurai, McHugh, Pollak, & Björgvinsson, 2020; Beard, Rifkin, Silverman, & Björgvinsson, 2019), and clinical judgement about a meaningful outcome for a low-intensity primary care intervention. Benchmarks included the following: at least 70% of participants complete at least six of the eight CBM sessions; at least moderate credibility (average score of five or greater on CEQ item 1); good treatment satisfaction (a CSQ-8 mean total score $\geq$ 24); and 50% of participants showing a decrease of 20% or more on HARS and GAD-7 scores.

Table 3. Open and randomized control trial benchmarks for feasibility and acceptability

	Open trial	Randomized control trial	
Benchmark	Mental habits	Mental habits	Symptom tracking
At least 70% of participants complete 6 or more of the 8 CBM sessions (%)	✓ 71.43% (10/14)	✗ 47.37% (9/19)	N/A
At least moderate credibility (average score of 5 or greater on the CEQ) *M* (*SD; N*)	✓ 7.21 (1.76; 14)	✓ 6.92 (2.10; 13)	✓ 5.79 (1.72; 19)
At least a CSQ-8 mean total score of 24 or greater *M* (*SD; N*)	✓ 29.69 (2.87; 13)	✓ 29.78 (1.86; 9)	✓ 27.82 (4.16; 17)
	ITT / PP	ITT / PP	ITT / PP
At least 50% of participants experience a 20% reduction in HARS scores (%)	✓ ✓ 11/14 (78.57%) 9/10 (90.00%)	✗ ✓ 5/19 (26.32%) 5/9 (55.66%)	✗ ✗ 7/21 (33.33%) 7/18 (38.89%)
At least 50% of participants experience a 20% reduction in GAD-7 scores (%)	✓ ✓ 10/14 (71.43%) 8/10 (80.00%)	✓ ✓ 10/19 (52.63%) 7/9 (77.78%)	✗ ✓ 10/21 (47.62%) 9/18 (50.00%)

Depiction of how the five benchmarks for treatment feasibility and acceptability were/were not met in the Open Trial and Randomized Control Trial (Mental Habits and Symptom Tracking). A checkmark indicates that the benchmark was met, while an 'X' indicates that it was not met. Under all checkmarks and 'Xs' are either the percentage of participants who met the benchmark, or the mean score and corresponding standard deviation.

ITT = intent-to-treat; PP = per protocol.

Feasibility and acceptability

Mental Habits met all feasibility and acceptability benchmarks. Ten of the 14 participants (71.43%) completed at least 6 CBM sessions. Scores on CEQ item 1 ranged from 4 to 9, with an average of 7.21 (*SD* = 1.76), indicating moderate-to-good treatment credibility. The mean total CSQ-8 score was 29.69 (*SD* = 2.87), indicating very high intervention satisfaction.

Symptoms

Given the small sample, we used last observation carried forward (LOCF) to handle missing follow-up data. We first compared symptom reduction against our benchmarks in the intent-to-treat sample (all participants enrolled in open trial). Of the 14 participants, 11 (78.6%) had a $\geq$ 20% decrease in HARS scores and 10 (71.4%) evidenced a decrease of 20% or greater in GAD-7 scores. We then compared obtained scores to benchmarks in the per-protocol sample (participants who completed at least six CBM sessions). Of the 10 participants, 9 (90%) had a $\geq$ 20% decrease in HARS scores and 8 (80%) evidenced a decrease of 20% or greater in GAD-7 scores. See Table 3.

Study 1 Discussion

Open trial results suggested that Mental Habits would be feasible to deliver and acceptable to primary care patients with anxiety disorders. 71% of participants completed all treatment sessions. This rate exceeds that of many other studies of computer-based / assisted treatment for anxiety in primary care (e.g., Bergman Nordgren et al., 2014; Newby et al., 2013). Interview and self-report measures of anxiety met our a priori benchmarks for effectiveness in both the intent-to-treat and per-protocol analyses. These initial findings suggested that a Stage IB feasibility and pilot RCT were warranted.

STUDY 2

We conducted a Stage 1B, small pilot RCT comparing Mental Habits to an active comparison, 'Symptom Tracking'. Eligibility criteria, screening and recruitment methods, and a priori benchmarks remained the same as in Study 1.

Study 2 Methods

Participants and setting

Initially, Study 2 participants were recruited from the same site as Study 1. During Study 2, the hospital at which the RI site was located closed. Due the long, chaotic closure process, many participants did not complete the study follow-up assessments, and we ultimately had to stop recruitment. As a result, we added a site in Massachusetts at an academically affiliated primary care clinic in an urban setting serving a racially, ethnically, and socio-economically diverse population. This practice had integrated behavioural health, staffed by four licensed social workers, a population health case manager, and a psychiatrist.

A total of 65 participants (RI = 49, MA = 16) consented to be screened. Ineligibility across sites (n = 22) was due to no primary diagnoses of GAD, SD, or PD/A (n = 9; 40.91%), lack of ability to read English (n = 1; 4.55%), cognitive impairment (n = 1; 4.55%), psychotic or other severe symptoms (n = 11; 50.00%). The remaining three who

did not enroll were not interested or lost to follow-up. Of the 65 screened, 40 participants (RI = 27, MA = 13) were eligible and enrolled (see Figure 1 for CONSORT Diagram).

Eligible individuals were randomized to Mental Habits (n = 19; RI = 12; MA = 7) or Symptom Tracking (*n* = 21; RI = 15; MA = 6). We used MinimPy software (Saghaei & Saghaei, 2011) for sequential allocation of participants to groups. Variables for randomization were scores on the Clinical Anxiety Scale, derived from the Hamilton Anxiety Scale (HARS-CAS; Snaith, Baugh, Clayden, Husain, & Sipple, 1982) and primary diagnoses (see Table 2). Groups did not significantly differ on gender or age (*ps* >.50). Participants were predominantly female (72.50%), and the mean age was 42 years (*SD* = 14.91) (see Table 2).

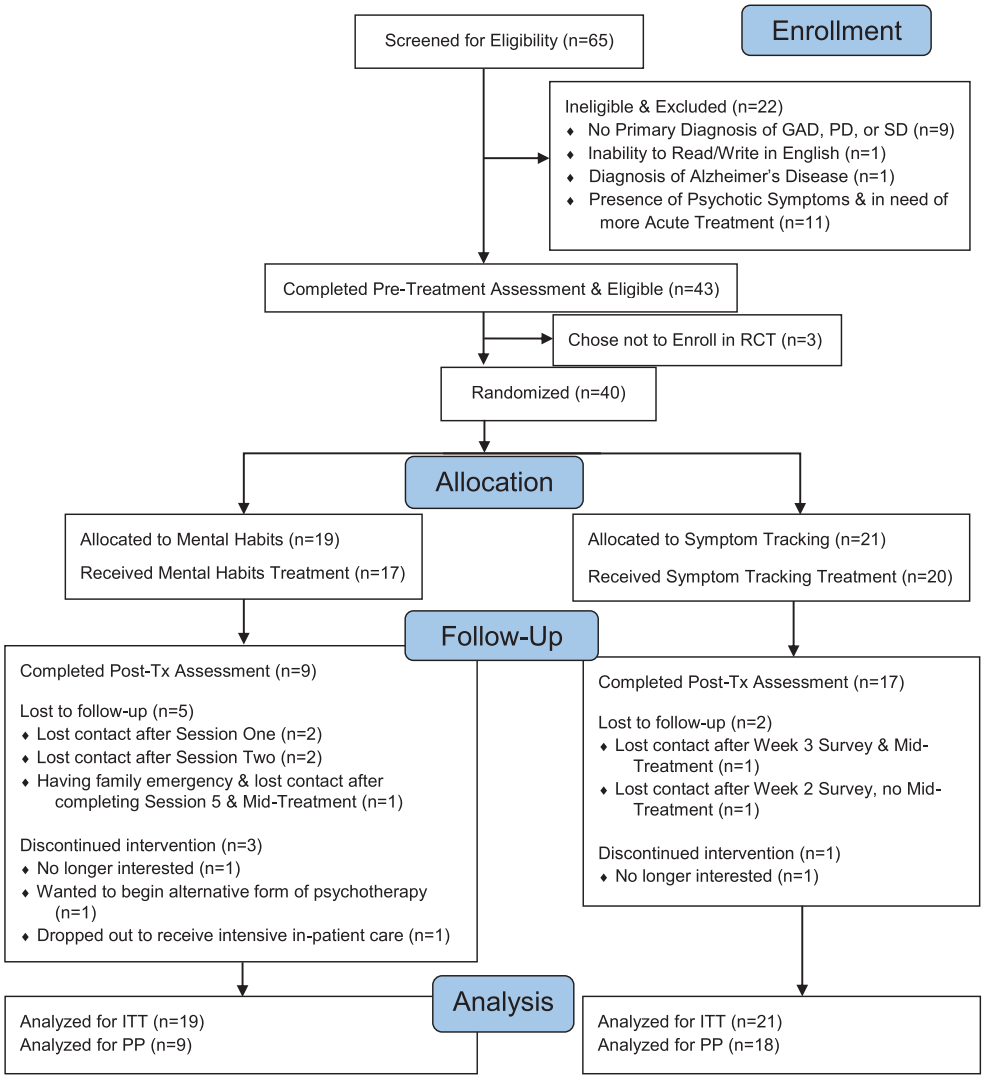

Figure 1. RCT CONSORT Diagram. The CONSORT diagram of participant flow in the Mental Habits and Symptom Tracking arms of the Randomized Control Trial. ITT = intent-to-treat; PP = per protocol.

Measures

All measures were the same as Study 1, except for the addition of cognitive bias measures to pilot the feasibility of studying potential mediators in a future, larger RCT. Cognitive bias was assessed via two computer tasks programmed using E-Prime 2.0 (Schneider, Eschman, & Zuccolotto, 2002).

Attention bias

To assess attention bias, we administered a version of the Posner spatial cueing task at each assessment point. In this task (based on Amir, Elias, Klumpp, & Przeworski, 2003), a word appeared on either the left or right side of the computer monitor in a rectangle. The word either validly cued the location of a probe that followed (probe appeared in same rectangle) or invalidly cued (probe appeared in opposite location's rectangle). Participants indicated which side of the screen the probe appeared. The task comprised 192 trials (66% valid, 33% invalid) and included 32 words each presented 6 times. Stimuli included eight words each for neutral, general threat, panic threat, and social threat. The spatial cueing task has demonstrated better reliability compared to other measures of attention bias (Hedge, Powell, & Sumner, 2018).

Interpretation bias

Interpretation bias was assessed with 120 trials of the WSAP (Beard & Amir, 2009). This assessment was similar to the CBM-I intervention, except that feedback about participants' responses was not provided, and stimuli were not personalized. We recorded the percentage of threat and benign interpretations endorsed. The WSAP is a widely used measure of interpretation bias with good internal consistency and test–retest reliability (Gonsalves et al., 2019).

Procedure: Symptom Tracking condition

All procedures were the same as Study 1, except for the addition of the Symptom Tracking condition. Participants assigned to Symptom Tracking followed the same procedures as those assigned to Mental Habits (e.g., in-person orientation session, two instructional videos, and weekly symptom surveys), with the exception that they did not complete ABM or CBM-I exercises, engage in weekly check-ins with the coach, and did not see online graphical feedback of their changes in GAD-7 and PHQ-9 scores over time. Symptom Tracking participants were each assigned a coach who met with them in person at the primary care practice for their first session, which included an orientation and their first surveys. After that, coaches sent subsequent weekly surveys to participants via SurveyMonkey (SurveyMonkey Inc., 2020) and monitored patients' scores. Just as in Mental Habits, coaches reached out to patients who evidenced clinical deterioration or suicidality and connected them with a licensed psychologist for a risk assessment.

Study 2 Results

Of the 40 participants enrolled, 65.00% ($n = 26$) completed a post-treatment assessment. A significantly greater percentage of participants completed the post-treatment in Symptom Tracking (80.95%; 17/21), than in Mental Habits (47.37%; 9/19), X^2 (1, 40) = 4.95, $p = 0.026$. There was also a difference in post-treatment assessment rates between

sites. At the MA site, 76.92% (10/13) of participants completed post-treatment vs 59.26% (16/27) at the RI site. One participant in the Mental Habits condition experienced clinical deterioration unrelated to study participation. This participant dropped out of the study to seek more intensive treatment after Session 1.

Feasibility and acceptability

In Mental Habits, 47% (9/19) of participants completed at least six intervention sessions. This rate was below our benchmark and notably lower than in Study 1. We thus explored potential site differences. At the MA site, 71% (5/7) completed at least six sessions, compared with only 33% (4/12) at the RI site. Treatment credibility ratings exceeded our benchmark for both conditions. Treatment satisfaction was also high in both arms as indicated by mean CSQ-8 scores and met our benchmark (see Table 3).

Attention bias

To pilot test the feasibility of assessing mediators of change in a future, adequately powered RCT, we examined changes in attention bias from pre- to the mid-treatment assessment (Mental Habits $n = 7$; Symptom Tracking $n = 11$). Consistent with prior studies (Amir et al., 2003), we eliminated response latencies for inaccurate trials (3% of trials) and response latencies < 50 ms or > 1700 ms (3.3% of trials). For each participant, we calculated the mean response latency for each word type and each cue condition and submitted these to a 2 (Condition: Mental Habits, Symptom Tracking) x 2 (Time: pre, mid-treatment) x 2 (Cue Type: valid, invalid) x 4 (Word Type: general threat, panic threat, social threat, neutral) analysis of variance (ANOVA) with repeated measurement on the last two factors. This revealed a main effect of validity, $F(1, 16) = 5.93$, $p = 0.027$, and word type, $F(1, 16) = 3.03$, $p = 0.038$. No other main effects or interactions were significant. We calculated an index of cue dependency (Amir et al., 2003) by subtracting response latencies for validly cued trials from the response latencies for invalidly cued trials. We submitted these scores to a 2 (Condition: Mental Habits, Symptom Tracking) x 2 (Time: pre, mid-treatment) x 2 (Word Type: general, panic, social, neutral). No effects were significant.

Interpretation bias

We piloted the feasibility of examining interpretation bias as a mediator. We examined changes from pre- to the mid-treatment assessment in 19 participants (Mental Habits n = 7; Symptom Tracking n = 12). We conducted a 2 (Condition: Mental Habits, Symptom Tracking) X 2 (Time: pre, mid) X 2 (Valence: benign, threat) repeated-measures ANOVA. There was a significant Condition X Time X Valence interaction, $F(1,17) = 7.69$, $p = .01$. We then examined the effect of condition separately for threat and benign interpretations. For threat interpretations, repeated-measures ANOVA revealed a significant interaction of time and condition, $F(1, 17) = 12.61$, $p = .002$. At pre-treatment, groups did not differ in threat interpretation, $t(17) = 0.49$, $p = .630$. By mid-treatment, Mental Habits had significantly better accuracy on threat trials (more rejections of threat) compared to Symptom Tracking, $t(17) = 3.58$, $p = .002$. Mental Habits participants' threat interpretations significantly improved from pre- to mid-treatment, $t(6) = 3.07$, $p = .022$; Symptom Tracking participants did not, $t(11) = 0.08$, $p = .940$. For benign interpretations,

there was no main effect of time, $F(1,17) = 3.16$, $p = .093$, or time by condition interaction, $F(1, 17) = 2.74$, $p = .116$.

Symptoms

Baseline HARS and GAD-7 scores did not differ between conditions (*ps* >.335). As the current RCT aimed to establish feasibility and acceptability, it was not designed to be powered to test between-group differences in symptoms. Thus, we focus here on comparing anxiety severity reduction to our a priori benchmarks and estimates of effect size. We used last observation carried forward (LOCF) to handle missing data.

Intent-to-treat

In Mental Habits, 26.32% (5/19) experienced a reduction in HARS of ≥ 20% compared to 33.33% (7/21) in Symptom Tracking. For Mental Habits, HARS scores reduced from pre ($M = 20.26$ ($SD = 5.98$); moderate severity) to LOCF ($M = 17.42$ ($SD = 8.03$); mild symptoms) with a medium effect size ($d = 0.531$). In Symptom Tracking, HARS scores also reduced from pre ($M = 18.52$ ($SD = 5.30$); moderate severity) to LOCF ($M = 15.19$ ($SD = 6.77$); mild symptoms) with a medium effect size ($d = 0.683$). At LOCF, 42.11% (8/19) of Mental Habits participants had HARS scores below the clinical cut-off (≤ 17; Antony, Orsillo, & Roemer, 2001) compared to 52.00% (11/21) in Symptom Tracking. To compare HARS scores between conditions, we calculated Hedges' g, and there was no evidence of a group difference ($g = 0.086$).

In the intent-to-treat sample, GAD-7 scores reduced by ≥ 20% in 52.63% (10/19) of participants in Mental Habits and 47.62% (10/21) of those in Symptom Tracking. For Mental Habits, GAD-7 scores reduced from pre ($M = 14.21$ ($SD = 4.33$); moderate severity) to LOCF ($M = 10.0$ ($SD = 6.14$); moderate symptoms) with a large effect size ($d = 1.149$). For Symptom Tracking, GAD-7 scores reduced from pre ($M = 14.10$ ($SD = 3.92$); moderate severity) to LOCF ($M = 10.43$ ($SD = 6.31$); moderate symptoms) with a large effect size ($d = 0.883$). At LOCF, 57.89% (11/19) in Mental Habits had GAD-7 scores below the cut-off of 10 compared to 57.14% (12/21) in Symptom Tracking. Comparing GAD-7 scores between conditions, there was no evidence of a group difference ($g = 0.129$).

Per protocol

Benchmarks for symptom reduction were met in the Mental Habits group, but not in Symptom Tracking. In Mental Habits, 55.56% (5/9) experienced a reduction in HARS from baseline of ≥ 20%, compared to 38.89% (7/18) in Symptom Tracking. For Mental Habits, HARS scores significantly reduced from pre ($M = 19.78$ ($SD = 6.38$); moderate severity) to LOCF ($M = 13.11$ ($SD = 8.58$); minimal symptoms) with a large effect size ($d = 1.032$). In Symptom Tracking, HARS scores reduced from pre ($M = 18.22$ ($SD = 5.64$); moderate severity) to LOCF ($M = 14.33$ ($SD = 6.93$); mild symptoms) with a medium effect size ($d = 0.76$). At LOCF, 66.67% (6/9) of Mental Habits participants had HARS scores below the clinical cut-off compared to 61.11% (11/18) in Symptom Tracking. Comparing LOCF HARS between conditions, there was a small-to-medium effect size favouring Mental Habits ($g = 0.457$).

GAD-7 was reduced by ≥ 20% in 77.78% (7/9) of participants in Mental Habits and 50.00% (9/18) of those in Symptom Tracking. For Mental Habits, GAD-7 scores

significantly reduced from pre (M = 12.56 (SD = 4.22); moderate severity) to LOCF (M = 6.89 (SD = 5.73); minimal symptoms) with a large effect size (d = 1.461). For Symptom Tracking, GAD-7 scores reduced from pre (M = 14.11 (SD = 3.80); moderate severity) to LOCF (M = 9.94 (SD = 6.34); minimal to moderate symptoms) with a large effect size (d = 0.966). At LOCF, 77.78% (7/9) in Mental Habits had GAD-7 scores below the cut-off of 10 compared to 61.11% (11/18) in Symptom Tracking. Comparing GAD-7 scores between conditions, there was a small effect favouring Mental Habits (g = 0.369).

Study 2 Discussion

In this pilot RCT, we compared Mental Habits to a Symptom Tracking condition. There were substantial site differences regarding treatment completion. While our treatment completion benchmark was met at the MA site, there was a high rate of dropout at the RI site, which led to notable differences in outcomes between intent-to-treat and per-protocol analyses. As noted above, the hospital at which the RI site was located underwent ownership changes and eventually closed during this study. Retention challenges at this site during Study 2 were likely largely due to the clinic's instability, as we did not experience these in Study 1 or at our MA site. However, in patients who completed the intervention protocol (at least 6 CBM sessions or at least 3 symptom tracking surveys), Mental Habits met all benchmarks for anxiety reduction with medium-to-large effect sizes, and was associated with improved interpretation bias (but not attention bias) outcomes.

General Discussion

We examined the feasibility, acceptability, and preliminary effectiveness of the first CBM intervention for transdiagnostic anxiety in primary care patients. Overall, results suggest that Mental Habits is a feasible and acceptable means of treating anxiety in primary care. As Study 1 and 2 were small pilot studies, we prioritized a priori benchmarks against which to compare our outcomes rather than conducting underpowered between-group tests. Feasibility was assessed by rates of treatment completion. Though fewer patients completed Mental Habits (47%) than our benchmark (70%) in Study 2, treatment completion exceeded the benchmark (71% in Study 1 and 71% at the MA site in Study 2) when Mental Habits was delivered at primary care sites with stable staffing and the expectation of a secure future patient relationship with the site. Further, though 47% treatment completion was below our benchmark, it is notable that we set a high standard for this outcome. In other studies of computer-based anxiety treatment in primary care, rates of treatment completion have not been reported (e.g., Graham et al. 2020) or have been between 32-41% (Bergman Nordgren et al., 2014; Newby et al., 2013). Acceptability of Mental Habits was high. Patients overwhelmingly rated the intervention as credible and reported high treatment satisfaction in both the open trial and RCT.

In both Study 1 and 2, Mental Habits exceeded our benchmarks for clinical effectiveness when delivered per protocol. In intent-to-treat analyses, the majority (52.63-71.43%) of patients assigned to Mental Habits self-reported clinically significant reductions in anxiety symptoms on the GAD-7 from pre- to post-treatment (LOCF). Interviewer-rated anxiety symptom reduction varied more between Study 1 and Study 2, with 78.57% of individuals assigned to Mental Habits in Study 1 evidencing a 20% or greater decrease in HARS score, compared to only 26.32% of Study 2 participants assigned to Mental Habits. It is of note that we had little study dropout in Study 1, and

only had one participant for whom in the intent-to-treat analyses we needed to carry forward the pre-treatment HARS score to the post-treatment score. In contrast, in Study 2, where our intent-to-treat (ITT) analyses of Mental Habits did not meet our a priori benchmark, study dropout required us to carry forward the pre-treatment assessment scores of nine participants and the mid-treatment assessment scores of two participants. Further, ITT analysis of HARS scores was more greatly impacted by study dropout than was ITT analyses of GAD-7 scores, as the HARS was only completed at pre, mid, and post-treatment, whereas the GAD-7 was completed weekly, enabling us the ability to carry forward more scores during the course of treatment for the GAD-7 than for the HARS.

Symptom Tracking was also found to be credible, satisfactory and potentially effective. As symptom monitoring is considered an active component of many cognitive-behavioural interventions for anxiety and has been found in other studies to reduce anxiety symptoms (e.g., Mavandadi, Benson, DiFilippo, Streim, & Oslin, 2015), this is not surprising. However, Mental Habits was generally associated with larger symptom improvement effect sizes and met more benchmarks for effectiveness.

Preliminary data also suggest that Mental Habits engaged its target of interpretation bias (reduction of threat interpretations). We did not find support for attention bias as a potential mediator of change. This is consistent with other studies and conclusions from a recent meta-analysis of CBM for anxiety (Fodor et al., 2020) in which CBM-I was found to have more reliable effects on anxiety than ABM. Although an adequately powered RCT is necessary to confirm target engagement (or lack thereof), future iterations of Mental Habits may be more efficient by only including the CBM-I component.

Interpretation of results is limited by the developmental nature of this work. Studies were designed with inadequate statistical power to detect group differences and thus, we relied on a priori benchmarks to provide a signal as to potential clinical effectiveness. Due to the instability of our RI site, there was a far higher rate of study dropout in Study 2, further limiting our ability to compare groups and to draw conclusions from intent-to-treat analyses. Mental Habits performed exceedingly well with regard to all benchmarks during Study 1 and the Study 2 per-protocol analyses.

We chose to develop an online, patient-delivered intervention, to reduce reliance on highly trained therapists and expand the reach of anxiety treatment to patients who have not previously had access. Consistent with other behavioural intervention technologies in primary care (e.g., Graham et al., 2020), we utilized non-professional coaches to enhance adherence to the intervention and monitor safety. In both study arms, participants had a coach who oriented them to the intervention and whom they knew was monitoring their progress and alerted to any clinical deterioration. However, due to the different nature of Mental Habits and Symptom Tracking, coaches spent more time with participants in Mental Habits. It is unclear how much this extra attention impacted outcomes. Future work is also needed to determine the benefit of coaches compared to a purely self-administered intervention.

We did not compensate participants for intervention sessions to obtain unbiased estimates of feasibility and acceptability. However, we did compensate for study assessments, and this payment could have influenced participants' satisfaction ratings. Though participants were excluded if they were already engaged in psychotherapy in the community or if they had any new initiation of pharmacotherapy or PCBH contact in the eight weeks prior to the study intake assessment, we did not monitor PCBH contacts during the study. It is unclear whether these may also have impacted results. Finally, we have limited information about Mental Habits in Latinx patients. Though the overall

racial–ethnic makeup of Study 2 was representative of the area population, all Latinx patients were randomized to Symptom Tracking. To our knowledge, the WSAP task used in Mental Habits has not yet been translated to Spanish (Gonsalves et al., 2019). Further research should examine the acceptability of Mental Habits in Latinx populations and as delivered in Spanish.

The need for low-intensity, scalable, anxiety treatments is great. When we began this work, we sought to develop an intervention which could easily and effectively be completed by the large number of primary care patients in need of anxiety treatment. As we write this manuscript, the world has been changed by the COVID-19 pandemic. There has been an urgent need to scale up telehealth. Mental Habits, in particular the personalized interpretation bias modification component, may offer an efficient and effective means of reaching patients with anxiety remotely and safely. A full-scale, hybrid implementation-effectiveness RCT is needed to extend this preliminary work.

Acknowledgements

This research was supported by the National Institute of Mental Health (MH097820) awarded to the first and last author and was registered on ClinicalTrials.gov (Identifier: NCT02579915). We are grateful to the entire staff and patients at the Family Care Center of Memorial Hospital of Rhode Island, especially Gretchen Sloane, Lisa Billington, Katalin Spremulli, Lauren Chan, Donna Haddock, Charles Eaton, and Justin Nash, and Brigham and Women's South Huntington Advanced Primary Care, especially Jane Erb, Sarah Gale, Debra Aponte, Janna Levi, and Leah Coburn, who made this work possible. We are grateful for all our clinical assessors, Sarah Garnaat, Marisa Sklar, Randy Whittles, Rachel Wiedemann, William Lechner, Maggie Evans, Miryam Yusufov, Kate Stewart, Jessica Strong, Bryana White, and Daniel Evans. We thank Danielle Shayani, Shawn Mozeika, Anna Schwartz, Spencer Price, Caroline Wingard, Monica Treviño, Janet Wanjiku, Veronica Fletcher, Sarah Fracasso Francis, Elizabeth Arriaga, Kate Silzer, Michelle Jolson, Esther Adegbulugbe, Michelle Menezes, and Molly Schulson for serving as coaches.

Conflict of interest

All authors declare no conflict of interest.

Author contributions

Risa B Weisberg (Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Writing – original draft; Writing – review & editing) Meghan A Gonsalves (Data curation; Formal analysis; Investigation; Project administration; Writing – original draft; Writing – review & editing) Ramya Ramadurai (Data curation; Formal analysis; Project administration; Writing – review & editing) Howard Braham (Methodology; Software; Writing – review & editing) Cara Fuchs (Data curation; Investigation; Methodology; Project administration; Supervision; Writing – review & editing) Courtney Beard (Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Writing – original draft; Writing – review & editing).

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

REFERENCES

Amir, N., Beard, C., Taylor, C. T., Klumpp, H., Elias, J., Burns, M., & Chen, X. (2009). Attention training in individuals with generalized social phobia: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *77*, 961. https://doi.org/10.1037/a0016685

Amir, N., Elias, J., Klumpp, H., & Przeworski, A. (2003). Attentional bias to threat in social phobia: Facilitated processing of threat or difficulty disengaging attention from threat? *Behaviour Research and Therapy*, *41*, 1325–1335. https://doi.org/10.1016/S0005-7967(03)00039-1

Amir, N., & Taylor, C. T. (2012). Interpretation training in individuals with generalized social anxiety disorder: a randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *80*, 497. https://doi.org/10.1037/a0026928

Amir, N., Weber, G., Beard, C., Bomyea, J., & Taylor, C. T. (2008). The effect of a single-session attention modification program on response to a public-speaking challenge in socially anxious individuals. *Journal of Abnormal Psychology*, *117*, 860–868. https://doi.org/10.1037/a0013445

Beard, C., & Amir, N. (2008). A multi-session interpretation modification program: Changes in interpretation and social anxiety symptoms. *Behaviour Research and Therapy*, *46*, 1135–1141. https://doi.org/10.1016/j.brat.2008.05.012

Beard, C., & Amir, N. (2009). Interpretation in social anxiety: When meaning precedes ambiguity. *Cognitive Therapy and Research*, *33*, 406–415. https://doi.org/10.1007/s10608-009-9235-0

Beard, C., Fuchs, C., Asnaani, A., Schulson, M., Schofield, C. A., Clerkin, E. M., & Weisberg, R. B. (2016). A pilot open trial of cognitive bias modification for panic disorder. *Cognitive Therapy and Research*, *40*, 792–798. https://doi.org/10.1007/s10608-016-9790-0

Beard, C., Ramadurai, R., McHugh, R. K., Pollak, J. P., & Björgvinsson, T. (2020). HabitWorks: Development of a CBM-I smartphone app to augment and extend acute treatment. *Behavior Therapy*. https://doi.org/10.1016/j.beth.2020.04.013

Beard, C., Rifkin, L. S., Silverman, A. L., & Björgvinsson, T. (2019). Translating CBM-I into real-world settings: Augmenting a CBT-based psychiatric hospital program. *Behavior Therapy*, *50*, 515–530.

Beard, C., Sawyer, A. T., & Hofmann, S. G. (2012). Efficacy of attention bias modification using threat and appetitive stimuli: A meta-analytic review. *Behavior Therapy*, *43*, 724–740. https://doi.org/10.1016/j.beth.2012.01.002

Beard, C., Weisberg, R. B., & Amir, N. (2011). Combined cognitive bias modification treatment for social anxiety disorder: a pilot trial. *Depression and Anxiety*, *28*, 981–988. https://doi.org/10.1002/da.20873

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

Bergman Nordgren, L., Hedman, E., Etienne, J., Bodin, J., Kadowaki, A., Eriksson, S., . . . Carlbring, P. (2014). Effectiveness and cost-effectiveness of individually tailored Internet-delivered cognitive behavior therapy for anxiety disorders in a primary care population: a randomized controlled trial. *Behaviour Research and Therapy*, *59*, 1–11. https://doi.org/10.1016/j.brat.2014.05.007

Bruss, G. S., Gruenberg, A. M., Goldstein, R. D., & Barber, J. P. (1994). Hamilton Anxiety Rating Scale Interview guide: joint interview and test-retest methods for interrater reliability. *Psychiatry Research*, *53*, 191–202. https://doi.org/10.1016/0165-1781(94)90110-4

Chavira, D. A., Golinelli, D., Sherbourne, C., Stein, M. B., Sullivan, G., Bystritsky, A., . . . Craske, M. (2014). Treatment engagement and response to CBT among Latinos with anxiety disorders in primary care. *Journal of Consulting and Clinical Psychology*, *82*, 392–403. https://doi.org/10.1037/a0036365

Cristea, I. A., Kok, R. N., & Cuijpers, P. (2015). Efficacy of cognitive bias modification interventions in anxiety and depression: meta-analysis. *The British Journal of Psychiatry*, *206*(1), 7–16. https://doi.org/10.1192/bjp.bp.114.146761

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, *31*, 73–86. https://doi.org/10.1016/S0005-7916(00)00012-4

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. (2002). *Structured clinical interview for DSM-IV-TR axis I disorders, research version* (patient ed.) (pp. 94–101). New York, NY: SCID-I/P.

Fodor, L. A., Georgescu, R., Cuijpers, P., Szamoskozi, Ş., David, D., Furukawa, T. A., & Cristea, I. A. (2020). Efficacy of cognitive bias modification interventions in anxiety and depressive disorders: a systematic review and network meta-analysis. *The Lancet Psychiatry*, *7*, 506–514. https://doi.org/10.1016/S2215-0366(20)30130-9

Gonsalves, M., Whittles, R. L., Weisberg, R. B., & Beard, C. (2019). A systematic review of the word sentence association paradigm (WSAP). *Journal of Behavior Therapy and Experimental Psychiatry*, *64*, 133–148. https://doi.org/10.1016/j.jbtep.2019.04.003

Goodwin, H., Yiend, J., & Hirsch, C. R. (2017). Generalized Anxiety Disorder, worry and attention to threat: A systematic review. *Clinical Psychology Review*, *54*, 107–122. https://doi.org/10.1016/j.cpr.2017.03.006

Graham, A. K., Greene, C. J., Kwasny, M. J., Kaiser, S. M., Lieponis, P., Powell, T., & Mohr, D. C. (2020). Coached mobile app platform for the treatment of depression and anxiety among primary care patients: A randomized clinical trial. *JAMA Psychiatry*, *77*, 906. https://doi.org/10.1001/jamapsychiatry.2020.1011

Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological Bulletin*, *137*, 940. https://doi.org/10.1037/a0024355

Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology*, *32*(1), 50–55. https://doi.org/10.1111/j.2044-8341.1959.tb00467.x

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*, 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Hirsch, C. R., Krahé, C., Whyte, J., Loizou, S., Bridge, L., Norton, S., & Mathews, A. (2018). Interpretation training to target repetitive negative thinking in generalized anxiety disorder and depression. *Journal of Consulting and Clinical Psychology*, *86*, 1017. https://doi.org/10.1037/ccp0000310

Hirsch, C. R., Meeten, F., Krahé, C., & Reeder, C. (2016). Resolving ambiguity in emotional disorders: The nature and role of interpretation biases. *Annual Review of Clinical Psychology*, *12*, 281–305. https://doi.org/10.1146/annurev-clinpsy-021815-093436

Kazdin, A. E., & Blase, S. L. (2011). Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspectives on Psychological Science*, *6*(1), 21–37. https://doi.org/10.1177/1745691610393527

Kimmel, R. J., Roy-Byne, P. P., & Cowley, D. S. (2014). Pharmacological treatments for panic disorder, generalized anxiety disorder, specific phobia, and social anxiety disorder. In P. Nathan & J. Gorman (Eds.), *A guide to treatments that work* (4th ed., pp. 463–506). New York, NY: Oxford University Press.

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

Larsen, D. L., Attkinson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: development of a general scale. *Evaluation and Program Planning*, *2*, 197–207. https://doi.org/10.1016/0149-7189(79)90094-6

Liebowitz, M. R. (1987). Social phobia. *Modern Problems of Pharmacopsychiatry*, *22*, 141–173. https://doi.org/10.1159/000414022

Linetzky, M., Pergamin-Hight, L., Pine, D. S., & Bar-Haim, Y. (2015). Quantitative evaluation of the clinical efficacy of attention bias modification treatment for anxiety disorders. *Depression and Anxiety*, *32*, 383–391. https://doi.org/10.1002/da.22344

Liu, H., Li, X., Han, B., & Liu, X. (2017). Effects of cognitive bias modification on social anxiety: A meta-analysis. *PLoS One*, *12*, e0175107. https://doi.org/10.1371/journal.pone.0175107

Mavandadi, S., Benson, A., DiFilippo, S., Streim, J. E., & Oslin, D. (2015). A telephone-based program to provide symptom monitoring alone vs symptom monitoring plus care management for late-life depression and anxiety: A randomized clinical trial. *JAMA Psychiatry*, *72*, 1211–1218. https://doi.org/10.1001/jamapsychiatry.2015.2157

McHugh, R. K., Whitton, S. W., Peckham, A. D., Welge, J. A., & Otto, M. W. (2013). Patient preference for psychological vs pharmacologic treatment of psychiatric disorders: a meta-analytic review. *The Journal of Clinical Psychiatry*, *74*, 595–602. https://doi.org/10.4088/JCP.12r07757

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the penn state worry questionnaire. *Behaviour Research and Therapy*, *28*, 487–495. https://doi.org/10.1016/0005-7967(90)90135-6

Mogoaşe, C., David, D., & Koster, E. H. (2014). Clinical efficacy of attentional bias modification procedures: An updated meta-analysis. *Journal of Clinical Psychology*, *70*, 1133–1157. https://doi.org/10.1002/jclp.22081

Mohlman, J. (2012). A community based survey of older adults' preferences for treatment of anxiety. *Psychology and Aging*, *27*, 1182. https://doi.org/10.1037/a0023126

Mulvaney-Day, T. M., Downey Piscopo, K., Korsen, N., Lynch, S., Karnell, L. H., Moran, G. E., … Ghose, S. S. (2018). Screening for behavioral health conditions in primary care settings: A systematic review of the literature. *Journal of General Internal Medicine*, *33*(3), 335–346. https://doi.org/10.1007/s11606-017-4181-0

Naim, R., Kivity, Y., Bar-Haim, Y., & Huppert, J. D. (2018). Attention and interpretation bias modification treatment for social anxiety disorder: A randomized clinical trial of efficacy and synergy. *Journal of Behavior Therapy and Experimental Psychiatry*, *59*, 19–30. https://doi.org/10.1016/j.jbtep.2017.10.006

Newby, J. M., Mackenzi, A., Williams, A. D., McIntyre, K., Watts, S., Wong, N., & Andrews, G. (2013). Internet cognitive behavioural therapy for mixed anxiety and depression: a randomized controlled trial and evidence of effectiveness in primary care. *Psychological Medicine*, *43*, 2635–2648. https://doi.org/10.1017/s0033291713000111

Ofonedu, M. E., Belcher, H. M. E., Budhathoki, C., & Gross, D. A. (2017). Understanding barriers to initial treatment engagement among underserved families seeking mental health services. *Journal of Child and Family Studies*, *26*, 863–876. https://doi.org/10.1007/s10826-016-0603-6

Onken, L. S., Carroll, K. M., Shoham, V., Cuthbert, B. N., & Riddle, M. (2014). Reenvisioning clinical science: Unifying the discipline to improve the public health. *Clinical Psychological Science*, *2* (1), 22–34. https://doi.org/10.1177/2167702613497932

Rounsaville, B. J., Carroll, K. M., & Onken, L. S. (2001). A stage model of behavioral therapies research: Getting started and moving on from stage I. *Clinical Psychology: Science and Practice*, *8*, 133–142. https://doi.org/10.1093/clipsy.8.2.133

Saghaei, M., & Saghaei, S. (2011). Implementation of an open source customizable minimization program for allocation of patients to parallel groups in clinical trials. *Journal of Biomedical Science and Engineering*, *4*, 734–739. https://doi.org/10.4236/jbise.2011.411090

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime (Version 2.0). [Computer software and manual]*. Pittsburgh, PA: Psychology Software Tools Inc..

Shear, M. K., Brown, T. A., Barlow, D. H., Money, R., Sholomskas, D. E., Woods, S. W., … Papp, L. A. (1997). Multicenter collaborative panic disorder severity scale. *American Journal of Psychiatry*, *154*, 1571–1575. https://doi.org/10.1176/ajp.154.11.1571

Snaith, R. P., Baugh, S. J., Clayden, A. D., Husain, A., & Sipple, M. A. (1982). The clinical anxiety scale: An instrument derived from the hamilton anxiety scale. *British Journal of Psychiatry*, *141*, 518–523. https://doi.org/10.1192/bjp.141.5.518

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, *166*, 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Stein, M. B., Sherbourne, C. D., Craske, M. G., Means-Christensen, A., Bystritsky, A., Katon, W., . . . Roy-Byrne, P. P. (2004). Quality of care for primary care patients with anxiety disorders. *American Journal of Psychiatry*, *161*, 2230–2237. https://doi.org/10.1176/appi.ajp.161.12.2230

SurveyMonkey Inc (2020). San Mateo, California, USA. www.surveymonkey.com

Thompson, V. L., Bazile, A., & Akbar, M. (2004). African Americans' perceptions of psychotherapy and psychotherapists. *Professional Psychological Research and Practice*, *35*(1), 19–26. https://doi.org/10.1037/0735-7028.35.1.19

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Research*, *168*, 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

van Dis, E. A., van Veen, S. C., Hagenaars, M. A., Batelaan, N. M., Bockting, C. L., van den Heuvel, R. M., . . . Engelhard, I. M. (2020). Long-term outcomes of cognitive behavioral therapy for anxiety-related disorders: a systematic review and meta-analysis. *JAMA Psychiatry*, 77, 265–273. https://doi.org/10.1001/jamapsychiatry.2019.3986

Verhaak, P. F., Prins, M. A., Spreeuwenberg, P., Draisma, S., van Balkom, T. J., Bensing, J. M., . . . Penninx, B. W. (2009). Receiving treatment for common mental disorders. *General Hospital Psychiatry*, *31*(1), 46–55. https://doi.org/10.1016/j.genhosppsych.2008.09.011

Watkins, E. R., Taylor, R. S., Byng, R., Baeyens, C., Read, R., Pearson, K., & Watson, L. (2012). Guided self-help concreteness training as an intervention for major depression in primary care: A Phase II randomized controlled trial. *Psychological Medicine*, *42*, 1359–1371. https://doi.org/10.1017/S0033291711002480

Weisberg, R. B., Beard, C., Moitra, E., Dyck, I., & Keller, M. B. (2014). Adequacy of treatment received by primary care patients with anxiety disorders. *Depression and Anxiety*, *31*, 443–450. https://doi.org/10.1002/da.22209

Yeung, E. S., & Sharpe, L. (2019). Cognitive Bias Modification for Social Anxiety: The Differential Impact of Modifying Attentional and/or Interpretation Bias. *Cognitive Therapy and Research*, *43*, 781–791. https://doi.org/10.1007/s10608-019-10012-3

Zhang, A., Borhneimer, L. A., Weaver, A., Franklin, C., Hai, A. H., Guz, S., & Shen, L. (2019). Cognitive behavioral therapy for primary care depression and anxiety: a secondary meta-analytic review using robust variance estimation in meta-regression. *Journal of Behavioral Medicine*, *42*, 1117–1141. https://doi.org/10.1007/s10865-019-00132-2

Received 12 August 2020; revised version received 30 November 2020