Mental Health on the Go: Effects of a Gamified Attention Bias Modification Mobile Application in Trait Anxious Adults

Tracy A. Dennis and
Hunter College, The City University of New York

Laura O’Toole
The Graduate Center, The City University of New York

Abstract

Interest in the use of mobile technology to deliver mental health services has grown in light of the economic and practical barriers to treatment. Yet, research on alternative delivery strategies that are more affordable, accessible, and engaging is in its infancy. Attention bias modification training (ABMT), has potential to reduce treatment barriers as a mobile intervention for stress and anxiety, but the degree to which ABMT can be embedded in a mobile gaming format and its potential for transfer of benefits is unknown. The present study examined effects of a gamified ABMT mobile application in highly trait anxious participants (N = 78). A single session of the active compared to placebo training reduced subjective anxiety and observed stress reactivity. Critically, the long (45 minutes) but not short (25 minutes) active training condition reduced the core cognitive process implicated in ABMT (threat bias) as measured by an untrained, gold-standard protocol.

Keywords

attention bias modification training; mobile application; anxiety; stress; gamification
This crisis in the mental health field has sparked a dialogue in psychology and other health-related fields about the need to develop alternative strategies for delivery of services (Christensen, Miller, & Muñoz, 1978; Harwood & L’Abate, 2010; Kazdin & Blase, 2011; Kazdin & Rabett, 2013; L’Abate, 2007; Mosa, Yoo, & Sheets, 2012; Rotheram-Borus, Swendeman, & Chorpita, 2012; Ryder, 1988). Particular attention has been paid to computerized and mobile interventions given their potential to serve as “disruptive innovations” that provide a qualitative leap in reducing cost and increasing accessibility of empirically-validated treatments (e.g., Barak, Hen, Boniel-Nissim, & Shapira, 2008; Kazdin & Rabett, 2013; Rotheram-Borus et al., 2012). Moreover, the ubiquity of mobile devices provides a unique opportunity to broaden the reach of psychological services to many individuals who might not otherwise have access (Dimeff, Paves, Skutch, & Woodcock, 2011; Kazdin & Rabbitt, 2013; Morris et al., 2010). Of the 91% of American adults who own cell phones, fully half of them, a number approaching 150,000,000, use mobile applications on their phones and 60% use their hand-held device to access the internet (Duggan, September, 2013).

The Games for Health movement (Buday, Baranowski, & Thompson, 2012; Ferguson, 2012; Kato, 2012; Rahmani & Boren, 2012) takes this idea further by exploring ways in which effective interventions can be translated into game format (e.g., game-like interfaces, points and rewards, and animated graphics) in order to reduce problems with treatment engagement and compliance and to increase use of prevention technologies. Although recent, widely-publicized research demonstrated that a video game successfully enhanced cognitive control in older adults (Anguera et al., 2013), this research is the exception. Thus, the promise of mobile technology for the delivery of mental health services remains vastly underexplored and understudied and it is largely unknown whether treatment approaches embedded in a game format – or gamification-result in transfer of benefits to gold-standard, lab-based assessments (e.g., Buday et al., 2012; Rahmani & Boren, 2012).

Attention bias modification training (ABMT) is an emerging computer-based therapeutic approach rooted in neurocognitive models of anxiety that overcomes many obstacles to treatment. It has been discussed as having significant potential both as an enhancement to state-of-the art psychological and pharmacological treatments for anxiety, as well as a stand-alone treatment (Bar-Haim, 2010; Hakamata et al., 2010). Thus, ABMT is a prime candidate for development into mobile, gamified interventions, but it is unknown whether ABMT effects will be retained in a gamified format. This is the goal of the present study.

ABMT emerged from research on core attention disruptions that to play a role in the etiology and maintenance of anxious pathology (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007; Brotman et al., 2007; Fox, Russo, Bowles, & Dutton, 2001; Fox, Russo, & Dutton, 2002; Mathews & Mackintosh, 1998; Mathews & MacLeod, 1985, 2002). In particular, exaggerated attention to threat, termed the threat bias, has been examined as a causal mechanism in anxiety (Hakamata et al., 2010; MacLeod, Rutherford, Campbell, Ebsworthy, & Holker, 2002; Mohlman, 2004; Puliafico & Kendall, 2006). The threat bias, thought to emerge in childhood (Puliafico & Kendall, 2006; Roy et al., 2008), is implicated in the emergence and expression of anxiety across diagnostic categories (Bar-Haim et al., 2007) and predicts a persistent course of anxiety from childhood to adulthood....
Bar-Haim et al., 2007; Pérez-Edgar et al., 2010; Pérez-Edgar et al., 2011). A meta-analysis (Bar-Haim et al., 2007) confirmed the presence of threat bias in both pediatric and adult anxiety and showed that this bias was robust across anxiety disorders but not present in non-anxious participants. This selective and exaggerated attention to threat may contribute to the continuity of anxiety by facilitating preferential processing of threat at the expense of pleasant cues or cues for safety. This in turn may spark a vicious cycle in which anxiety is heightened, attention to threat is further facilitated, and opportunities for disconfirmation of fear beliefs are minimized (e.g., Hofmann, 2007).

In ABMT, attention is trained away from threat by creating attentional competition between a threat and non-threat stimulus and repeatedly directly participants’ attention towards the non-threat stimulus. A recent meta-analysis of ABMT using this simple technique showed that ABMT not only resulted in reduced threat bias with a large effect size ($d = 1.16$), but also produced significantly greater reductions in anxiety than PT with a medium effect size ($d = .61$); these effects that were consistently sustained at follow-ups varying between one to six months (Hakamata et al., 2010). In addition to reductions in state anxiety for highly trait anxious participants (Eldar & Bar-Haim, 2010), several studies have shown that ABMT results in reduced symptoms of several anxiety disorders (generalized anxiety disorder: Amir, Beard, Burns, & Bomyea, 2009; social phobia: Amir, Beard, Taylor, et al., 2009; pathological worry: Hazen, Vasey, & Schmidt, 2009; social anxiety disorder: Klumpp & Amir, 2010) comparable to the effect size of a typical 12-session CBT intervention after one month (Hakamata et al., 2010). Stress reactivity is also reduced by ABMT. For example, a study with high anxious youth showed that participants in an ABMT versus control training condition showed less stress-induced anxiety following a task-based stressor (Bar-Haim, Morag, & Glickman, 2011).

Given this evidence base that ABMT has the potential to effectively reduce threat bias, anxiety symptoms and anxiety-related stress reactivity, combined with the fact that ABMT overcomes many treatment barriers by being brief and inexpensive, ABMT is a prime candidate for mobile-based intervention approaches. Consistent with this, several studies have modified ABMT for use on mobile devices or to be delivered via the internet (Amir & Taylor, 2012; Boettcher et al., 2013; Carlbring et al., 2012; Enock & McNally, 2013; MacLeod, Soong, Rutherford, & Campbell, 2007). However, these studies resulted in inconsistent results and, crucially, did not address the need to increase individual engagement and acceptability of interventions (Buday et al., 2012; Ferguson, 2012; Kato, 2012; Rahmani & Boren, 2012). For youth and young adults in particular, gamifying interventions can make them more appealing, reduce stigma, and increase compliance. Moreover, given the effects of ABMT on stress reduction, the wider use of ABMT to reduce stress and stress-related disorders would be greatly facilitated by the development of enjoyable, gamified interventions that are also readily accessible and affordable by being mobile.

The goal of the present study was to examine effects of a gamified ABMT mobile application in a sample of high trait anxious participants. To do so, we took the core components of the gold-standard ABMT protocol (the dot probe task; MacLeod, Mathews, & Tata, 1986) and designed an appealing game around the basic task-parameters while
incorporating video game-like features such as animated characters, points, and sound effects (See Figure 1). We administered a “short training” condition of the app (25 minutes with 20 minutes of rest) and a “long training” condition of the app (45 minutes with no rest) based on commonly-used numbers of trials for ABMT trials (Eldar et al., 2012; Eldar & Bar-Haim, 2010) as an initial exploration of “dosage effects.” We examined whether this gamified ABMT app affected the threat bias, anxiety, and stress reactivity of trait anxious individuals in ways similar to traditional, lab-based ABMT by testing the following hypotheses: **Hypothesis 1**: Threat bias measured via an independent computerized task will be reduced among participants in the ABMT versus placebo training (PT) version of the app. **Hypothesis 2**: Subjective anxiety and observed anxious stress reactivity will be reduced among participants in the ABMT versus PT version of the app. **Exploratory Hypothesis 3**: Because threat bias is the hypothesized target of ABMT, we explored whether reductions in the independent measure of threat bias predicted reduced subjective anxiety and stress reactivity, and whether this varied between the short and long training conditions.

**Method**

**Participants**

Participants were adults recruited from an Introduction to Psychology course at an urban university in New York City. The long training condition included 38 participants (27 females; ages 17–50, $M_{age} = 22.34, SD = 6.91$) who were randomly assigned to either an ABMT or PT condition (18 per group). The short training condition was conducted after the long training condition was completed. It included 40 participants (28 females; ages 17–38, $M_{age} = 20.23, SD = 4.08$) who were randomly assigned to either an ABMT or PT condition (20 per group). However, two of these participants were excluded after completing the study due to failure to complete the app, leaving 18 in the attention training group and 20 in the PT condition. Demographic characteristics for each of the four groups are reported in Table 1.

**Procedure**

The procedure was identical for both the long and short training conditions. Participants were recruited if they scored +1 standard deviation above the mean for college students on trait anxiety (a score of 49) using the State-Trait Anxiety Inventory (STAI; Spielberger, 1983). Participants spent approximately 2 h in the laboratory. After a brief questionnaire period during which they completed demographic questions and self-report of state anxiety and depressed mood, participants were seated 65 cm from a 17 in monitor to complete the pre-training threat bias assessment using the dot probe. Following this, participants sat comfortably at a table in a separate room and completed the ABMT or PT on an iPod Touch (fourth generation). After the training session, participants immediately reported on their state anxiety and completed the post-training threat bias assessment using the dot probe. Following this, participants reported on positive and negative mood (Profile of Mood States questionnaire (POMS; McNair, Lorr, Heuchert, & Droppleman, 2003), completed the Trier Social Stress Task (TSST; Kirschbaum, Pirke, & Hellhammer, 1993), and again reported on positive and negative mood using the POMS.
Measures

**Baseline mood questionnaires**—Baseline trait anxiety and depression symptoms for each of the four groups are reported at the bottom of Table 1. The trait anxiety score was derived from the STAI; scores range from 20 to 80 with higher scores indicating greater anxiety (Spielberger, 1983). Depression symptoms were measured using the Beck Depression Inventory II (BDI-II; Beck, Steer, & Brown, 1996). Scores range from 0 to 63: 0–13 is minimal, 14–19 is mild, 20–28 is moderate, and 29–63 is severe.

**The dot probe task and stimuli**—The dot probe task is most commonly used computerized task to measure the threat bias and best reflects naturalistic conditions such that threatening and non-threatening stimuli compete for attention, rather than threatening stimuli being presented alone (Bar-Haim et al., 2007; Mathews & Mackintosh, 1998). The dot probe protocol followed parameters consistent with the Tel-Aviv University/National Institute of Mental Health protocol (http://www.tau.ac.il/~yair1/ABMT.html). Stimuli for the dot probe task are pictures of 20 different individuals (10 males, 10 females) from the NimStim stimulus set (Tottenham et al., 2009) with one female taken from the Matsumoto and Ekman set (1989). Stimuli were programmed using E-Prime version 2.0 (Schneider, Eschman, & Zuccolotto, 2002).

During each trial a pair of the task, two pictures are presented, either angry-neutral face pairs or neutral-neutral face pairs (depicted by the same individual). The pictures are shown above and below a fixation cross, with 14 mm between them. The task includes 120 trials (80 angry-neutral and 40 neutral-neutral). Each trial, as depicted in Figure 1, is comprised of: (a) 500 ms fixation, (b) 500 ms face-pair cue, (c) target (probe) in the location of one of the faces until a response is made via the left or right mouse button to indicate the direction in which the arrow is pointing, and (d) 500 ms ITI. Participants were asked to respond as quickly and as accurately as possible whether the arrow was pointing to the left or the right. Probes were equally likely to appear on the top or bottom, in the location of the angry or neutral face cues, and pointing to the left or the right.

Reaction times (RT) were filtered by removing responses that were faster than -3SD from an individual's mean and slower than +3SD from an individual's mean. Using correct trials only, the dot probe task yields three threat bias scores (attentional bias, vigilance, disengagement) by comparing reaction times (RT) following the different probe conditions. The angry face probe is when the target replaces the angry face from a pair of angry and neutral faces, while the neutral face probe is when the target replaces the neutral face from a pair of angry and neutral faces. A baseline probe is when the target replaces either face from a pair of two neutral faces.

The threat bias score is calculated as the average RTs for neutral probes–angry probes. Higher scores indicate an attentional bias toward threatening information, such that participants respond faster when the probe appears in the location of the angry face versus the neutral face, i.e., under conditions in which threat competes for attention with non-threat. This bias can be driven by the speed of attentional capture by threat (vigilance) or the length of attentional hold by threat (disengagement). The vigilance score is calculated as the average RTs for baseline probes–angry probes. Higher scores indicate greater attentional
capture by threat, such that participants respond faster when the probe appears in the location of the angry face versus when no threatening face is presented. The disengagement score is calculated as the average RTs for neutral probes–baseline probes. Higher scores indicate greater attentional hold by threat, such that participants are faster to respond when no threat is presented versus when they have to disengage and shift attention to the location of the neutral face.

**ABMT and PT conditions of the app**—Participants were assigned either to the ABMT or PT condition of the app. Both experimenters and participants were blind to whether assignment was to the ABMT or PT condition, and participants were blind as to the purposes of the app. Participants sat comfortably at a table, and were given an iPod Touch. They were instructed by the experimenter, "In this game two animated characters will appear on the screen. Shortly after, they will burrow into a hole. One of them will cause a path of grass to rustle behind it. With your finger, trace the path of the rustling grass, beginning from the burrow. Try to complete this task as quickly and as accurately as possible.” Then, they were given one practice screen in which the swiping motion on the touch screen was demonstrated and they were able to practice. The screen did not advance to the game until the swiping was correctly executed. Experimenters remained to answer any questions about the game. After each block of trials (40 trials) experimenters recorded the accrual of points and end-of-round feedback (see below).

For every trial, two cartoon characters (sprites), one showing an angry expression and one showing a neutral/mildly positive expression, appeared simultaneously on the screen for 500 ms (see Figure 2, left side). Next, both sprites simultaneously “burrowed” into the grass field. In the ABMT condition, a trial of grass appeared in the location of the non-threat character for every trial, whereas in the PT condition, a trial was equally likely to appear in the location of the angry or neutral sprite. The grass remained until the participants responded (see Figure 2, right side). Paths were divided into separate “tufts” of grass (randomly varying between five and eight tufts), and when a tuft was correctly traced it was illuminated.

Points were accrued based on speed and accuracy (see Appendix for scoring details). Participants were given feedback after each trial by being presented with one of three possible fixation “jewels” varying in color and accompanying sound: a red jewel with the lowest pitch sound indicated slower responses/less accuracy; a purple jewel with a medium pitch sound indicated moderate speed/accuracy; and a gold jewel with a higher pitched sound indicated the fastest speed/accuracy (see Appendix for detailed algorithm for determining feedback jewel). When errors were made (e.g., swiping the grass path towards rather than away from the burrow hole; not touching any portion of the grass path), a feedback sound was given (a high pitched “huh?”). Points were accrued on every trial because game play would not advance without a correct response.

The short training condition included consisted of 12 blocks of 40 trials for a total of 480 trials (25 minutes of game play with 20 minutes of breaks given consisting of a 10-minute break after round 160 and a 10-minute break after round 320). The long training condition consisted of 16 blocks of 40 trials for a total of 640 trials (45 minutes of game play with
brief breaks given as needed). The total duration of experimental time (game play and breaks, equaling 45 minutes) was the same across both the short and long training conditions. The number of short training trials were selected based on ABMT studies showing significant reductions in threat bias, anxiety, and stress reactivity following a single session of 480 trials (Eldar & Bar-Haim, 2010; Klumpp & Amir, 2009) and multiple sessions of 480 trials (Eldar et al., 2012; Li, Tan, Qian, & Liu, 2008). Thus, the long training condition provided a number of training trials that exceeded the number that have been associated with significant effects of ABMT on anxiety- and stress-related outcomes in traditional ABMT studies.

Pre- and post-training state anxiety—The state anxiety score was derived from the STAI; scores range from 20 to 80 with higher scores indicating greater anxiety (Spielberger, 1983). State anxiety was measured immediately prior to the app training and immediately after completing the app training.

Stress reactivity—Following the app training and post-training threat bias assessment, the Trier Social Stress Test (TSST) was administered (Kirschbaum et al., 1993). The TSST included both a social-evaluative threat (giving a speech for three minutes after three minutes of preparation) and a lack of control task (three minute arithmetic task). Both tasks were video-recorded and completed in front of two researchers who were described as judges. Participants were informed that their performance would be compared to others in the study and that a voice-frequency analysis and analysis of non-verbal behaviors would be conducted. The TSST was not administered prior to attention training because acute stress may induce shifts in threat-related attention on the dot probe task (Bar-Haim et al., 2010) thus distorting the measurement of pre-training bias.

Anxious behavior coding: Behaviors were coded as present or absent during the speech and mental arithmetic stressors in 10 second time bins. Behaviors consisted of flight behaviors from Troisi (1999): looking down/away from the judge; closing the eyes; drawing the chin in toward the chest; crouching; being still or freezing. Additionally, nervous speech (e.g., “umm” or “hmm”) and verbal expressions of frustration (e.g., “Oh my goodness!” profanities, or groaning) were also coded. Videos were coded by four research assistants; for 15 of the videos, two research assistants completed the coding in order to assess reliability. Reliability (α = .78) was calculated using Krippendorff’s alpha for nominal data (present/absent). These three sub-scores (anxious behavior, nervous speech, verbal expressions of frustration) were analyzed separately and were summed into (a) a total score and (b) a total nervous speech score comprised of instances of nervous speech and verbal expressions of frustration. These two summary scores were arithmetic sum of all instances of these codes across the entire coding period (both speech and arithmetic task).

Self-report of mood: Self-reported mood was recorded before and after the TSST using the 65-item Profile of Mood States (POMS; McNair et al., 2003). Participants are instructed to indicate on a five-point scale how well each adjective describes their current mood (not at all to extremely). The POMS measures six different mood states (Tension/Anxiety, Depression/
Dejection, Anger/Hostility, Vigor/Activity (reverse scored), Fatigue/Inertia, Confusion/Bewilderment) which are combined to generate a Total Negative Mood score.

Analytic Approach

To test the central hypotheses that the ABMT condition of the app would reduce threat bias, subjective anxiety, and stress reactivity relative to the PT condition, and to explore “dosage” effects (short versus long training), a series of 4 (Training Condition: short training ABMT, short training PT, long training ABMT, long training PT) X 2 (Gender) analyses of covariance (ANCOVAs) were conducted. The ANCOVA approach allowed us to analyze for group differences in threat bias, subjective mood, and stress reactivity after the training procedure (post-training assessment) controlling for differences in these measures at baseline (pre-training assessment or pre-TSST mood). This analytic approach has greater power than a repeated measures ANOVA when assessing treatment effects between randomly assigned groups (Van Breukelen, 2006).

Dependent variables were: post-training threat bias, post-training vigilance, post-training disengagement, post-training state anxiety, post-stressor negative mood, and anxious behaviors and nervous speech during the stressor. Covariates were: pre-training threat bias measures, pre-training state anxiety, and pre-stressor negative mood. Behavior and speech during the TSST did not have a baseline measure to be used as a covariate.

While almost half the sample was 18–19 years of age, the age of the other half of the sample ranged from 20–50. Thus, we examined whether baseline mood and attention bias scores differed across the broad age groups: 18–19 year-olds (n = 35), adults in their 20’s (n = 23), and adults 30 years of age and above (n = 8). Because sample sizes were uneven, we conducted analyses using non-parametric independent samples Kruskal-Wallis tests. Several baseline measures differed across age groups (trait anxiety, vigilance, and threat bias, all \( p \)'s < .05 with the oldest participants showing the lowest scores on these three variables) and thus age in years was included as a covariate in all analyses below.

One participant was missing baseline threat bias data due to too many error responses (long training attention training condition) and seven participants refused to participate in the TSST (three short training PT, one short training ABMT, one long training PT, two long training ABMT). These participants were excluded from analyses on the effects of Training Condition on these outcomes (n = 1 excluded from threat bias analyses and n = 7 excluded from analyses with stress reactivity measures obtained during the TSST).

Results

The four Training Condition groups (short training ABMT, short training PT, long training ABMT, long training PT) did not differ on any demographic variables or in baseline anxious and depressed mood (ethnicity/race \( p > .07 \), all other \( p \)'s > .32; see Table 1). The means and standard deviations of all pre- and post-training attention bias scores, self-reported anxiety, and stress reactivity measures (self-reported mood before and after the stressor, anxious behaviors and nervous speech during the stressor) are presented in Table 2. In addition, we
Effects of Training Condition on Attention Bias

There was a significant main effect of Training Condition on threat bias \([F(3,65) = 3.15, p = .03, \eta^2 = .13]\) and disengagement \([F(3,65) = 4.94, p = .004, \eta^2 = .19]\). The effect for vigilance did not reach significance \([F(3,65) = 0.33, p = .80, \eta^2 = .01]\). When controlling for pre-training bias score and age, participants in the long training ABMT condition \((M = -10.37, SE = 5.42, 95\% CI = -21.19 – 0.45)\) showed less biased attention to threat after training relative to the long training PT condition \((M = 6.99, SE = 4.56, 95\% CI = -2.11 – 16.09)\) and relative to the short training ABMT condition \((M = 6.86, SE = 4.61, 95\% CI = -2.35 – 16.07; \text{both } p’s < .02)\). Effects for disengagement mirrored those for threat bias: participants in the long training ABMT condition \((M = -6.00, SE = 3.87, 95\% CI = -13.73 – 1.73)\) showed less difficulty disengaging from threat after training relative to the long training PT condition \((M = 5.25, SE = 3.26, 95\% CI = -1.25 – 11.75)\) and relative to the short training ABMT condition \((M = 8.58, SE = 3.29, 95\% CI = 2.02 – 15.140)\). Moreover, counter to prediction, participants in the short training ABMT condition showed greater difficulty disengaging from threat relative to short training PT condition \((M = -7.04, SE = 3.81, 95\% CI = -14.66 – 0.57, p = .01)\).

Effects of Training Condition on Anxious Mood and Stress Reactivity

Figure 3 presents participants’ state anxiety at the pre- and post-assessments. The main effect of Training Condition \([F(3,66) = 2.48, p = .07, \eta^2 = .10]\) showed that, as predicted, controlled for pre-training subjective anxiety, participants in the ABMT condition reported reduced subjective anxiety immediately after training relative to PT condition, but only for the short training conditions \((M = 35.10, SE = 1.40, 95\% CI = 32.31 – 37.88 \text{ versus } M = 39.42, SE = 1.61, 95\% CI = 36.22 – 42.63, p = .045)\). In addition, the short training ABMT condition resulted in reduced anxiety after training relative to the long training ABMT condition \((M = 39.19, SE = 1.64, 95\% CI = 35.91 – 42.47, p = .06)\) and relative to the long training PT condition \((M = 39.93, SE = 1.36, 95\% CI = 37.22 – 42.64, p = .02)\).

Effects of Reductions in Threat Bias on Anxiety and Stress Reactivity

To test exploratory hypotheses that reductions in threat bias would predict reductions in anxiety and stress reactivity, we conducted a series of hierarchical multiple regressions in which age, gender, baseline mood measures (pre-app state anxiety or pre-stressor mood), and baseline attention bias score (either threat bias, vigilance, or disengagement score prior to attention training) were entered in the first step; change in the corresponding attention bias measure from pre-to post-training was entered in the second step; training condition (ABMT or PT) was entered in the third step; and the interaction between training condition and change in attention bias was entered in the fourth step. Analyses were conducted separately for the short and long trainings in order to allow direct comparison between ABMT and PT conditions\(^1\).

**Short training**—Tests for moderation did not reach significance. However, regardless of training condition, reductions in threat bias \([\beta = -0.52, F(1, 32) = 6.00, p = .02, \Delta R^2 = .11]\).
and reductions in difficulty disengaging from threat \( \beta = -0.54, F(1, 32) = 4.52, p = .04, \Delta R^2 = .09 \) were associated with lower negative mood after the TSST. In addition, reductions in difficulty disengaging from threat was associated with less nervous speech expressed during the TSST \( \beta = -0.69, F(1, 29) = 6.20, p = .02, \Delta R^2 = .15 \). Figure 4 depicts a plot of these associations adjusted for the contribution of baseline attention bias, age, and gender.

**Long training**—As predicted, changes in threat bias from pre-to post-training moderated the effects of Training Condition \( \beta = -1.07, F(1, 26) = 3.14, p = .04, 1\text{-}tailed, \Delta R^2 = .08 \). Specifically, as depicted in Figure 5, we found that those in the ABMT condition showed reduced negative mood after the TSST, but only if they also showed decreased vigilance following app play [only the decrease vigilance line was significantly different from zero: \( t(33) = 1.90, p = .03, 1\text{-}tailed \)].

**Discussion**

The present study provides evidence that an alternative delivery strategy for ABMT – a gamified mobile app – shows transfer of benefits to independent, untrained lab-based measures of anxiety and stress reactivity after a single session of training. First, the app generalized to changes in the gold-standard measure of threat bias using the dot probe task: the long training ABMT resulted in reduced threat bias and difficulty disengaging relative to PT and relative to the short training ABMT. In addition, both long and short training versions of ABMT resulted in reductions in subjective and observed anxiety and stress reactivity. Taken together, results suggest that even after only a single session of play, this ABMT app may reduce and possibly prevent acute stress responses, and therefore act as, what some have termed, a “cognitive vaccine” (see also Holmes et al., 2009 and Browning et al., 2012 in reference to depressed mood). The efficacy of more extended, frequent, and briefer sessions of app use to reduce anxiety and stress, neurocognitive mechanisms underlying treatment effects, and efficacy of the app when used outside the lab are crucial directions for future research.

In the short training condition, ABMT relative to PT resulted in reduced subjective state anxiety, but, counter to predictions, was also associated with greater difficulty disengaging from threat measured via the dot probe. Although termed “short” training, 25 minutes of training (480 trials) is consistent with the duration of previous studies showing efficacy of ABMT to reduce anxiety/stress in a single session (Eldar & Bar-Haim, 2010; Klumpp & Amir, 2009) or multiple sessions (Eldar et al., 2012; Li et al., 2008). The significant effect of short training ABMT on subjective anxiety reflects that 25 minutes may represent an adequate amount of training to result in improved mood. However, because short training ABMT was also associated with greater difficulty disengaging from threat – rather than the

---

1We also tested whether training condition on anxiety and stress reactivity were mediated by changes in attention biases by conducting a series of hierarchical multiple regressions using the PROCESS macro for SPSS (Hayes, 2013). This procedure tests the product of the coefficient for (\( \alpha \)) the effect of the independent variable (training condition) to the mediator (change in threat bias, vigilance, or disengagement) with (\( \beta \)) the effect of the mediator to the dependent variable (anxiety and stress reactivity) when the independent variable is taken into account. Corresponding pre-training measures for the dependent variable, along with age, gender, and baseline mood measures were entered as covariates. The 95% confidence interval for the direct path (\( \alpha \beta \)) was calculated. There were no significant mediation effects.
increases in attentional control that other studies have documented (Amir, Beard, Taylor, et al., 2009; Amir, Weber, Beard, Bomyea, & Taylor, 2008) – this finding suggests the influence of other potential mechanisms in the alleviation of anxious mood, such as exposure and habituation (Beard, 2011). This question is an important future direction for the ABMT research field as a whole, especially given recent lab- and internet-based trials suggesting equal treatment effects for ABMT compared to placebo conditions (e.g., Carlbring et al., 2012), and even compared to conditions that train attention towards threat (e.g., Boettcher et al., 2013).

Relevant to this, for the short training, regardless of whether it was ABMT or PT, participants who showed decreases in threat bias and difficulty disengaging from threat showed less stress reactivity (observed anxious speech during the stressor and negative mood after the stressor). One possible explanation is that mere exposure to threat stimuli when they compete with non-threat stimuli can result in reductions in behavioral threat bias for some, a change which has anxiolytic and stress reduction effects (Beard, 2011; Boettcher et al., 2013; Carlbring et al., 2012). Additionally, the structured 10-minute breaks during the short version of the app training may have contributed to the reduction of stress reactivity by giving participants time to consolidate effects of the training (Abend, Karni, Sadeh, Fox, & Pine, 2013).

In contrast to findings for the short training ABMT, for participants given the long training condition of the app, ABMT relative to PT was associated with predicted reductions in threat bias and difficulty disengaging. These findings are consistent with previous studies showing reductions in threat bias and greater attentional control (ability to disengage from threat) following traditional ABMT (Hakamata et al., 2010). The massed presentation of trials in the long versus short versions of the training (i.e., without substantial breaks in the long version) may have contributed to these differential effects on the ability to disengage from threat. That is, massed presentation may be necessary for an acute shift in threat bias performance over a single session. It is unclear whether massed presentation over multiple sessions is necessary to induce reductions in threat bias, or whether shorter, more frequent sessions also can effectively reduce threat bias. This question is particularly important given that mobile applications are typically used in brief spurts (Purcell, November 2011).

Results also showed that subjective mood was influenced in interaction with changes in vigilance: the long version of ABMT resulted in reduced negative mood during the stressor, but only for those participants showing decreases in vigilance after app play. This suggests that changes in relatively automatic attentional capture, rather than later attentional hold, may have a more direct influence on whether ABMT results in reductions in stress reactivity. Research that uses temporally-sensitive measures such as electroencephalography and scalp-recorded event-related brain potentials may be particularly well-suited to delineate the potential impact of early/automatic versus later/controlled neurocognitive changes on ABMT efficacy (Eldar & Bar-Haim, 2010; O'Toole & Dennis, 2012; Smith, Cacioppo, Larsen, & Charttrand, 2003). Indeed, a recent study suggests that treatment-related reductions in early attention capture (the N1 ERP component) influence the impact of ABMT on stress reactivity (O’Toole & Dennis, under review). Such methodological additions may address growing concern about the context-sensitivity of reaction-time based...
measures of threat bias, which can reduce measurement reliability (Brown et al., 2013).
Indeed, in the present study, participants showed variability in baseline measures of threat bias, although there were no significant group differences, thus making it difficult to fully understand the degree to which the long-version of ABMT reduced threat bias. Larger sample sizes in future research will help increase confidence that variability at the baseline in both threat bias measures and measures of other individual differences (e.g., depression) is not influencing results. In addition, the development and validation of additional measures of threat bias is a crucial future direction for the field as a whole.

Given their potential to profoundly reduce barriers to mental health services, it is important to clarify whether this and other kinds of ABMT apps can be conceptualized as “cognitive vaccines” that prevent or disrupt the trajectory affective psychopathology (Browning et al., 2012). In this context, the concept of cognitive vaccine suggests the ability to recalibrate habits of attention that, in the case of the threat bias, rigidly tune the individual towards threat and away from safety cues. This recalibration likely expands the type of information available for decision making and coping (Stout, Shackman, & Larson, 2013; Todd, Cunningham, Anderson, & Thompson, 2012) and thus serves to increase cognitive, affective and behavioral flexibility. Greater flexibility may in turn disrupt the vicious cycle of anxiety in which threat bias heightens anxious arousal during stress, attention to threat is further facilitated, coping strategies become increasingly inflexible, and opportunities for disconfirmation of fear beliefs are minimized (Hofmann, 2007). If this flexibility can be maintained, it could both ameliorate anxiety severity and stress reactivity and bootstrap the efficacy of other treatments such as cognitive behavioral therapy (Bechor et al., in press; Rapee et al., 2013). The impact of cognitive bias modification protocols on the efficacy of other treatments and on other affective-cognitive processes (Abend et al., 2013) is a crucial direction for future research. These questions highlight the exciting potential of mobile and gamified treatment approaches, particularly in a preventive or adjunctive treatment context.

Taken together, these results support the utility of an alternative delivery approach for ABMT and provide initial insights into “dosage” effects of a single session. There are, however, several methodological limitations that should be considered when interpreting findings. First, participants in the present study were trait anxious but not clinically anxious. While including a normatively anxious sample is a reasonable first research step and is highly relevant to questions about stress and anxiety prevention, it provides no direct evidence about whether the app is effective in clinical groups. Second, unlike traditional ABMT protocols that use multiple threat-related and neutral stimuli (whether words, faces, or complex emotional pictures), the current app used only one threat and one non-threat cartoon character. This design may limit generalizability of the app, although some of the most robust findings of the present study were that the long training condition of the ABMT condition resulted in reduced threat bias measured via the dot probe, thus supporting the generalizability of the app to a novel an untrained context. Third, it remains unclear whether use of the app will be effective in non-lab contexts, such as in the daily life of the individual, where adherence cannot be as easily supervised (e.g., Carlbring et al., 2012; Enock & McNally, 2013).
While the mobile format of the app presents methodological challenges, it simultaneously provides the tremendous benefit of instant access so that training can be completed anywhere (e.g., before attending a stressful event) with a low barrier to entry (a device with Apple iOS) and in a form that is unobtrusive. Future research should examine both acceptability (willingness to use the app and adherence to treatment) and efficacy of long-term app use. Moreover, the longevity of effects and the optimal number of trials and sessions requires focused research attention.

Results showed that a single session of gamified ABMT effectively reduces threat bias, subjective anxiety, and stress reactivity. Findings add to the growing body of research illustrating that evidence-based treatment mechanisms can be embedded into mobile (Amir & Taylor, 2012; Enock & McNally, 2013; Holmes et al., 2009) and gamified protocols, including those that target a range of cognitive biases such as interpretation biases (Hallion & Ruscio, 2011; Mobini, Reynolds, & Mackintosh, 2012). These technologies are crucial targets for future research and hold promise in preventative, treatment, and self-help contexts.

**Acknowledgments**

We are grateful to Lee Dunn for her support and insight throughout the research project, and to Sam Kaufman and Shawn Feeney for their outstanding efforts on development of the mobile application.

This research was supported by grants from the National Institutes of Mental Health (NIMH) K01MH075764 and S06GM060654 awarded to Tracy A. Dennis. This research was also made possible by Grant RR03037 form the National Center for Research Resources (NCRR), a component of the National Institutes of Health.

**Appendix**

The total awarded score for a stage of play (40 rounds) is the **sum of two scores: the speed and accuracy scores**. If either the speed or accuracy score is a rejection, the total score is a rejection.

**Speed Score**

This is a simple map-to-a-points value from the delay between the end of the “jewel” feedback animation and the moment the first finger touched the screen ($d$). Speed score is never a rejection.

18750 max 0.01, min 1, 1-d3

**Accuracy Score**

There are two paths of interest, the drawn and target grass paths. (The grass path for the mean/untargeted face is ignored.) Each path is abstracted to three features:

- the centers of the bounding boxes,
- the widths of the bounding boxes, WD and WT, and
- the heights of the bounding boxes, HD and HT.
We next calculate some properties between the pair of drawn and target paths:

- the distance between the two centers divided by 480, \( CD \),
- the **absolute** difference between the drawn and target bounding box widths, divided by 480, \( XD = \text{ABS}(WD - WT) / 480 \),
- the **absolute** difference between the drawn and target bounding box heights, divided by 480, \( YD = \text{ABS}(HD - HT) / 480 \).

Now, we consider the accuracy score a rejection if any of the following are true:

- The signs (directions) of the drawn and target paths differ on either axis. That is, if a finger went left and the path went right.
- If the width of the drawn path’s bounding box is less than 10% that of the target path’s with, or vice versa with regard to the height.

If no rejection condition, the accuracy score is: \( 16250 \cdot 1 - CDD + XD + YD \)

**End-of-Round Feedback**

The feedback at the end of the round (i.e., a trial) is determined by the round-level score.

<table>
<thead>
<tr>
<th>Score</th>
<th>Jewel</th>
<th>Sound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rejection</td>
<td>None</td>
<td>error noise</td>
</tr>
<tr>
<td>Under 21,000 pts</td>
<td>Red</td>
<td>low-pitch positive noise</td>
</tr>
<tr>
<td>21,000 through 24,999</td>
<td>Purple</td>
<td>mid-pitch positive noise</td>
</tr>
<tr>
<td>25,000 or higher</td>
<td>Gold</td>
<td>high-pitch positive noise</td>
</tr>
</tbody>
</table>

**References**


*Clin Psychol Sci.* Author manuscript; available in PMC 2015 September 01.


Beck, AT.; Steer, RA.; Brown, GK. Manual for Beck Depression Inventory II (BDI-II). San Antonio, TX: Psychology Corporation; 1996.


Clin Psychol Sci. Author manuscript; available in PMC 2015 September 01.


Mathews A, MacLeod C. Induced processing biases have causal effects on anxiety. Cognition and Emotion. 2002; 16(3):331–354.


Purcell K. Half of adult cell phone owners have apps on their phones. Pew Research Center. 2011 Nov.


Van Breukelen GJ. ANCOVA versus change from baseline: more power in randomized studies, more bias in nonrandomized studies. Journal of Clinical Epidemiology. 2006; 59(9):920–925. [PubMed: 16895814]

Clin Psychol Sci. Author manuscript; available in PMC 2015 September 01.
Figure 1.
The dot probe task.
Figure 2.
Screen shots of the app game play.
Figure 3.
State anxiety before and after training.
The graph shows a scatter plot with points representing the relationship between the reduction in threat bias from pre to post and post TSST negative mood (POMS). The line of best fit indicates a negative correlation, with an $R^2$ value of 0.158 for linear regression.
The diagram illustrates the relationship between the reduction in difficulty disengaging from pre to post and the post TSST negative mood (POMS). The regression line shows a negative correlation, with an $R^2$ value of 0.124 indicating that 12.4% of the variance in post TSST negative mood can be explained by the reduction in difficulty disengaging from pre to post.
Figure 4.
For short training version of both the ABMT and PT conditions, pre- to post-training reductions in threat bias and in difficulty disengaging from threat were associated with lower levels of negative mood after the TSST (top row). Pre- to post-training reductions in difficulty disengaging from threat were also associated with less nervous speech during the TSST (bottom row).
Figure 5.
For the long training version only, the association between training condition and post-TSST negative mood was moderated by changes in vigilance.
## Table 1

Participant Demographics, Trait Anxiety, and Depression Symptoms

<table>
<thead>
<tr>
<th>Training Condition</th>
<th>Short Training</th>
<th>Long Training</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ABMT</td>
<td>PT (n = 20)</td>
</tr>
<tr>
<td>Gender (% women)</td>
<td>60</td>
<td>80</td>
</tr>
<tr>
<td>Age (years)</td>
<td>20.65 (5.11)</td>
<td>19.8 (2.78)</td>
</tr>
<tr>
<td>Education (years)</td>
<td>14.55 (2.05)</td>
<td>15.18 (3.71)</td>
</tr>
<tr>
<td>Ethnicity (frequency)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>American Indian</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Asian</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Native Hawaiian or Other Pacific Islander</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Black or African American</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>White</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>More than one race</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Trait Anxiety</td>
<td>48.44 (7.85)</td>
<td>51.25 (9.91)</td>
</tr>
<tr>
<td>Depression Symptoms</td>
<td>15.65 (9.32)</td>
<td>15.20 (10.20)</td>
</tr>
</tbody>
</table>

*Note.* Standard deviations are in parentheses. ABMT = Attention bias modification training; PT = placebo training.
### Table 2

Pre- and Post-Assessment Anxiety, Depression Symptoms, Mood, and Behavior Scores for each Training Condition

<table>
<thead>
<tr>
<th></th>
<th>Pre-Assessment</th>
<th>Post-Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Short Training</td>
<td>Long Training</td>
</tr>
<tr>
<td></td>
<td>ABMT</td>
<td>PT</td>
</tr>
<tr>
<td>Threat Bias</td>
<td>0.97 ± (15.38)</td>
<td>4.93 ± (27.38)</td>
</tr>
<tr>
<td>Vigilance</td>
<td>-4.22 ± (18.64)</td>
<td>1.65 ± (24.53)</td>
</tr>
<tr>
<td>Disengagement</td>
<td>5.22 ± (17.98)</td>
<td>3.40 ± (16.86)</td>
</tr>
<tr>
<td>State Anxiety</td>
<td>36.33 ± (11.64)</td>
<td>38.80 ± (10.49)</td>
</tr>
<tr>
<td>Mood (Pre TSST)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mood (Post TSST)</td>
<td>77.28 ± (24.09)</td>
<td>76.95 ± (34.39)</td>
</tr>
<tr>
<td>Anxious Behaviors</td>
<td>94.44 ± (18.08)</td>
<td>82.83 ± (19.76)</td>
</tr>
<tr>
<td>Nervous Speech</td>
<td>13.88 ± (5.58)</td>
<td>10.50 ± (5.45)</td>
</tr>
</tbody>
</table>

Note. Standard deviations presented in parentheses. ABMT = Attention bias modification training; PT = placebo training.

The TSST occurred at post-assessment only. Main effects of Training Condition are indicated as follows:

* Long ABMT < Long PT and Short ABMT;

* Short ABMT > Short PT.

† Short ABMT < Short PT, Long ABMT, Long PT.