Negative Global Metacognitive Biases are Associated with Depressive and Posttraumatic Stress Disorder Symptoms and Improve with Cognitive Training

Sam Agnoli[1,3], Henry Mahncke[6], Sarah-Jane Grant[6], William P. Milberg[3,5], Michael Esterman[1,2,3,4], Joseph DeGutis[1,3,5]

[1]Boston Attention and Learning Lab, VA Boston Health Care System, Boston, MA
[2]National Center for PTSD, VA Boston Health Care System, Boston, MA
[3]Translational Research Center for TBI and Stress Disorders (TRACTS), VA Boston Health Care System, Boston, MA
[4]Department of Psychiatry, Boston University Chobanian and Avedisian School of Medicine, Boston, MA
[5]Department of Psychiatry, Harvard Medical School, Boston, MA
[6]Posit Science Corporation, San Francisco, CA, USA

Corresponding Author and Contact for Data Access:

Joseph DeGutis

VA Boston Health Care System

150 South Huntington Ave (182-JP),

Boston, Massachusetts, 02130

Email: degutis@wjh.harvard.edu

Abstract

Metacognition, thinking about thinking, is disrupted in several clinical populations. One aspect of metacognition, global metacognitive bias (difference between objective and self-reported abilities), has shown to be particularly relevant to clinical functioning. However, previous studies of global metacognitive biases in clinical populations have not measured objective and self-reported abilities relative to normative samples, making quantification of the severity of biases difficult. Additionally, few studies have examined whether cognitive interventions can improve metacognitive biases and none have examined how this relates to depressive/PTSD symptoms. In 84 participants with mild traumatic brain injury (77% Veterans), a population whose self-reported cognitive deficits are often worse than their objective deficits, we assessed PTSD/depressive symptoms and self-reported and objective measures of global cognition. We used age-adjusted norm-based z-scores for cognition measures and calculated bias by subtracting objective from self-report scores. Participants then received 13 weeks of either targeted cognitive training or entertainment games training, with both conditions providing performance feedback. Participants were measured at baseline, immediately post-training, and 3 months post-training. We found large negative metacognitive biases in those with depression (z-score difference=-1.77), PTSD (-1.47), and depression+PTSD (-2.29). Notably, metacognitive bias improved after both targeted and entertainment training and was strongly associated with depressive and PTSD symptom improvements (r=.-41/-.42, respectively). These effects endured after 3-months of no contact. These findings show that depression/PTSD are associated with substantial negative global metacognition biases and provide initial evidence that cognitive training can improve biases and depressive/PTSD symptoms.

Keywords: global metacognition, depression, post-traumatic stress disorder, cognitive training

Metacognition, the capacity to evaluate and regulate one's thinking (Dunlosky and Metcalfe, 2008), has received increased attention for its role in mental health (Seow et al., 2021). An essential aspect of metacognition is awareness of one's performance, traits, and abilities (see Fleming and Lau, 2014). One component of metacognitive awareness is sensitivity, i.e., accuracy, or how closely self-evaluations are related to objective performance (Nelson, 1984). In contrast, metacognitive bias refers to one's calibration, i.e., tendency for self-evaluations to be either above vs. below objective performance, as measured by the difference between self-reported and objective abilities, traits, or performance on a particular task. Positive metacognitive biases are quite common (e.g., better-than-average effect, Alicke and Govorun, 2005) and are associated with increased life satisfaction and positive affect (Dufner et al., 2019). In contrast, clinical disorders including depression and anxiety have often been associated with more negative metacognitive biases (underconfidence) compared to healthy controls in 'local' trial-by-trial tasks (e.g., perceptual decision making, Rouault et al., 2022; for a review, see Hoven et al., 2019). Recent work has highlighted that more 'global' (i.e., across a domain) metacognitive biases in cognitive ability and health are associated with depression/PTSD symptoms (Agnoli et al., 2023). Compared to local measures, global metacognitive measures may be more functionally and clinically relevant (Seow et al., 2021) in addition to affording greater reliability (Benwell et al., 2022). However, studies have yet to compare global metacognitive biases between normative samples and clinical groups (e.g., depression and PTSD) to better characterize the extent of metacognitive dysfunction, which was a goal of the current study. Additionally, the current study aimed to measure potential changes in global metacognitive biases in cognition after 3 months of cognitive training, and how changes are related to clinical symptom improvements.

Several studies have observed depression-related negative metacognitive biases when participants make task-specific performance judgments, i.e., local metacognition, including perceptual ability (Rouault et al., 2023), social and nonsocial knowledge judgements (Fu et al., 2005), and verbal memory (Soderstrom et al., 2011). Far fewer studies have examined local metacognitive awareness in PTSD, though the existing studies show negative biases similar to depression (e.g., Sacher et al., 2018). These biases in local domains inform more global self-performance estimates (Rouault et al., 2019), and evidence suggests that depressive symptoms are also associated with more negative global metacognitive biases (in perception and self-esteem, Hoven et al., 2022). Agnoli and colleagues (2023) further showed that depressive and PTSD symptoms were associated with global metacognitive biases in both cognition and health (Agnoli et al., 2023). One model posits that depressed moods make it easier to retrieve negative information about oneself, leading to poorer self-reported abilities and more negative metacognitive biases, and potentially to more depressive thoughts and behaviors (e.g., negative schemas and avoidance, Hoven et al., 2023; Manos et al., 2009). Important limitations of the global metacognitive awareness literature are that previous studies have either inadequately measured self-reported cognition (by using a 6-item scale, Agnoli et al., 2023) or indirectly indexed abilities by using a general self-esteem measure (e.g., “I am able to do things as well as other people”, Hoven et al., 2022). As a result, studies have yet to examine how global self-reported and objective abilities relate to normative samples in order to precisely measure the magnitude of dysfunction in metacognitive bias (i.e., using z-scores derived from normative populations). The current study addressed these issues by measuring subjective and objective global cognition using validated, normed measures to better quantify the degree of miscalibration in those with PTSD and depression.

In addition to the degree of metacognitive bias in clinical populations, how metacognition changes over time and how this relates to clinical symptoms also remains poorly characterized. Agnoli et al. (2023) found that, in a veteran sample, changes in global metacognitive bias over a two-year period were significantly related to depression/PTSD symptom changes (ρ=-.25/-.33, respectively). Fox and colleagues (2023) showed that 12 weeks of cognitive behavioral therapy (CBT) reduced both anxious-depression symptoms and local metacognitive bias on a perception task, with a significant but modest correlation between these measures (r=-.12). While CBT overtly challenges negative cognitive distortions related to clinical symptoms (e.g., overgeneralizing upon isolated failures, Beck 1970), cognitive training with intensive feedback may be an alternative method to reduce global metacognitive biases. By not specifically targeting clinical symptoms, cognitive training could also provide mechanistic support for whether reducing metacognitive biases can drive clinical symptom improvements (a causal mechanism suggested by Capobianco et al., 2019).

The current literature provides some indirect support for cognitive training improving global metacognitive biases and clinical symptoms. ‘Targeted' games (e.g., focusing on training specific processes) and entertainment games (e.g., Tetris, first-person shooter games) have shown to improve objective and self-reported cognition (Alvarez et al., 2008, Calkins et al., 2015; for a meta-analysis, see Motter et al., 2016, though see Naismith et al., 2011), as well as some improvements in depression and anxiety (for a review, see Fleming et al., 2017, for limitations, see Abd-Alrazaq et al., 2022). However, it remains unclear if either objective cognitive improvements, changes in subjective cognition/metacognitive biases, or both are related to clinical symptom improvements. One trial in adults with major depression and cognitive dysfunction observed improvements in mood symptoms after a month of cognitive

control training (Gunning et al., 2021). These mood improvements were associated with improvements in both self-reported and objective cognitive measures. However, another study found that while objective cognitive performance improved after 10 weeks of training targeting memory, attention, and executive functioning, it was not related to improvements in depressive symptoms (Elgamal et al., 2007).

It could be that feedback from targeted and entertainment cognitive training may help to recalibrate metacognitive biases (decrease negative biases; Katyal et al., 2023), informing more global metacognition (Rouault et al., 2019) and leading to improvements in depressive thoughts and behaviors. In their seminal review/synthesis of the video game literature, Granic and colleagues (2014) argue that the immediate, adaptive feedback in entertainment games (or gamified targeted cognitive training) rewards continuous engagement (Lyons, 2015) and teaches individuals about their self-efficacy in and out of the games (Pavlas et al., 2010; Dos Santos et al., 2016). Further, this feedback may increase perceived user benefits of games (Hassan et al., 2019). In the present study, we sought to compare whether targeted cognitive training vs. active control entertainment games improves objective cognition and global metacognitive bias, and whether these potential improvements are related to changes in depressive/PTSD symptoms.

The Present Study

The present study sought to address three goals: 1) To adequately characterize the magnitude of global metacognitive biases in a clinical sample in relation to normative populations; 2) To test whether depression/PTSD are associated with global negative metacognitive bias, replicating our previous study (Agnoli et al., 2023) with a more comprehensive measure of self-reported cognition; and 3) To determine if targeted and/or entertaining cognitive training can change metacognitive bias, and if this is correlated with

changes in depressive/PTSD symptoms. To answer these questions, we analyzed an existing dataset from a randomized control trial (*n*=84) of mostly Veterans (77%) with mild traumatic brain injury (mTBI) across the country with self-reported and/or objective cognitive impairment (Mahncke et al., 2021), as we previously found more negative metacognitive biases in a similar veteran population (Agnoli et al., 2023). Veterans with mTBI are a good population to study negative metacognitive biases given that 72% report moderate to severe cognitive impairment (Seal et al., 2016) despite only 35% meeting DSM-5 criteria for mild objective cognitive dysfunction (Riley et al., 2019). Participants completed 13 weeks of either targeted cognitive training or an active control of entertainment games. All participants were administered the Cognitive Failures Questionnaire (CFQ; Broadbent et al., 1982) which globally indexes self-reported cognitive mistakes/weaknesses and has an adequate range in normative populations (Goodman et al,. 2022), as well as a global battery of objective computer-based cognitive tasks spanning attention, working memory, and executive functioning. By contrasting normed-based z-scores on both the CFQ and objective cognitive battery, we could better quantify the degree of global metacognitive biases. Additionally, because self-reported and objective cognitive ability were measured across timepoints before, immediately after, and 3-months after the experimental training or active control condition, we could examine how changes in metacognitive bias were related to changes in depression/PTSD and how they may endure. To our knowledge, this is the first study to characterize global metacognitive biases in relation to normative populations as well as the first to examine how cognitive training impacts global metacognitive biases and their relation to depression/PTSD.

Methods

Participants

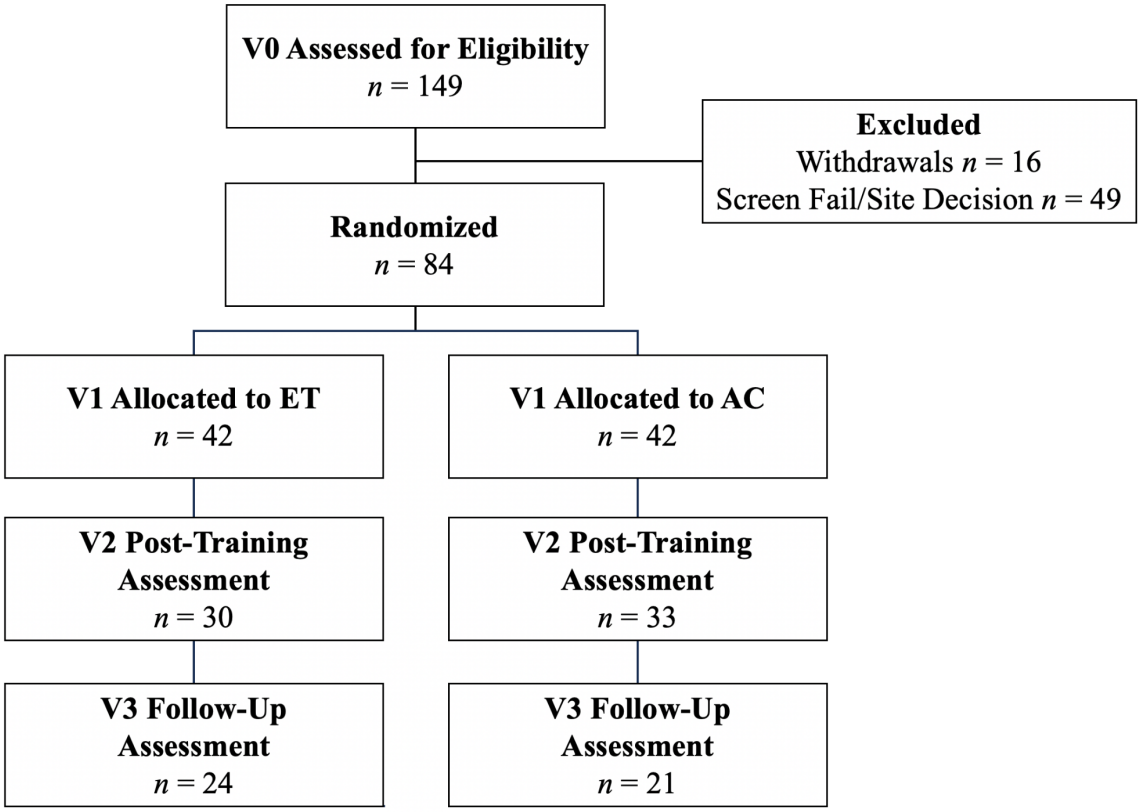
Participants were part of the BRAVE trial, a multi-site, randomized clinical trial of cognitive training (Mahncke et al., 2021). Participants were recruited from five military and VA sites (Walter Reed National Military Medical Center, Schofield Barracks, VA Boston Healthcare System, Michael E. DeBakey VA Medical Center, and VA Connecticut Healthcare System). Inclusion and exclusion criteria were chosen to identify participants with (i) a history of mTBI; and (ii) evidence of current cognitive impairment based on neuropsychological criterion ≥1 standard deviation (SD) below the norm based on a cognitive battery from the Automated Neuropsychological Assessment Metrics (ANAM; Reeves et al., 2006) , or (iii) self-reported cognitive impairment ≥2 SD below the norm based on the Ruff Neurobehavioral Inventory (RNBI) attention, executive functioning, learning, or memory subscales (Young et al., 2009).

Exclusionary criteria included a history of penetrating head wounds or a diagnosis of moderate/severe TBI, in-patient status, a diagnosis with cognitive consequences (e.g., schizophrenia, bipolar disorder, cancer, multiple sclerosis; however, common mTBI comorbidities including PTSD, depression, and chronic pain, were not exclusion criteria), or participation in a concurrent clinical trial that could influence data collection (participation in standard treatments, e.g., antidepressants, were not exclusion criteria). Participants with significant visual impairment were excluded, as were those with active suicidal ideation/behavior. A total of 84 participants aged 18-50 (M=32.61, SD=8.74) were included in the sample, which was 82% male, 79% White, and 77% Veteran or active duty military members. For a diagram of CONSORT flow before, during, and after the cognitive training intervention, see Figure 1. This study was approved by the VA Boston Healthcare System institutional review board, the coordinating center, and each trial site. All participants provided written informed consent, and the study was carried out in accordance with the declaration of

Helsinki.

Figure 1.

Diagram of CONSORT Flow



Note: A complete CONSORT flow is shown in Fig. 1. ET refers to targeted cognitive training. AC refers to active control. Drop/withdrawal rates were not significantly different between groups (p=0.554, chi-square), and there were no significant differences between completers and non-completers (data not shown) nor between the experimental treatment drop/withdraw and active control drop/withdraw groups (data not shown). Reasons for drop/withdrawal were typically the time commitment of study participation or change in life circumstances. Number of sessions completed was not significantly different between groups (experimental treatment 38.7±24.4, active control 42.4±23.4, p=0.470).

Clinical Measures

The Beck Depression Inventory (BDI-II) is a reliable, well-validated 21-item self-report questionnaire assessing the severity of depressive symptomatology (Beck, Steer, and Carbin, 1988). Cutoff scores of ≥19 indicate probable depression (Homaifar et al., 2016). The PTSD Checklist Civilian (PCL-C) is a reliable, well-validated 20-item self-report measure of PTSD, corresponding to DSM-V symptoms for PTSD (Weathers et al., 2013; Blevins et al., 2015). Cutoff scores of ≥40 indicate probable PTSD (Karstoft et al., 2014).

Self-Reported Cognition, Objective Cognition, and Metacognitive Bias

The current study included 3 different self-reported measures of cognitive functioning. The most comprehensive of which, the Cognitive Failures Questionnaire (CFQ; Broadbent et al., 1982), was prioritized in analyses. The CFQ is a 25-item measure of cognitive errors experienced in daily life, including forgetfulness, distractibility, and thinking blunders. In this study, we asked participants about their cognitive failures in the past month. Recent work has suggested that the CFQ is particularly applicable across a global range of abilities rather than specific cognitive domains (Goodman et al., 2022). For more information on how age-adjusted normative data for the CFQ and other cognitive measures were obtained, see the Supplementary Materials. The RNBI is a 21-item questionnaire on a 4-point Likert scale that asks participants to compare their current daily functioning to their premorbid condition (Young et al., 2009). The RNBI has four cognitive scales: attention and concentration, executive functions, learning and memory, and speed and language. A normed T-score of <70 (recommended by the RNBI manual as documenting significant post-morbid impairment) on any of these four cognitive scales was required for a participant to be included based on the RNBI (self-report) criterion.

Our objective cognitive battery had 8 validated tasks. The Rey Auditory Learning Test

(RAVLT) sum of trials 1-5 and delayed recall were used as measures of learning and memory, respectively (Schmidt, 1996). To measure working memory, we included Digit Span (sum of forwards, backwards, and sequencing trials, Wechsler Adult Intelligence Scale, Wechsler, 2008) and Symbol Span from the Wechsler Memory Scale (Wechler, 1997). Finally, as measures of executive function, we used the flanker and set-shifting task from the Executive Abilities: Measures and Instruments for Neurobehavioral Evaluation and Research (EXAMINER; Kramer et al., 2014). Our objective cognition composite score was a summation of the mean total scores of all 8 tasks[1]. All tasks were norm and age-adjusted based on data in healthy populations (see Supplementary Materials). In addition, the Automated Neuropsychological Assessment Metrics (ANAM) was completed during visit 0 as the objective deficits screening measure. The ANAM is a neuropsychological composite measure that generally focuses on processing speed (Reeves et al., 2006). An ANAM TBI Battery score $\geq$1 was used to characterize cognitive impairment and subsequent inclusion given our sample's history of mTBI (see Mahncke et al., 2021). Both self-reported and objective screening measures were only used to identify inclusion criteria and were not used to define self-reported/objective/metacognitive bias scores. Though this data set was originally presented in Mahncke et al. (2021), because measures were norm-adjusted and used a different cognitive composite, the current study represents distinct results than those presented in Mahncke et al.

To measure global metacognitive bias, we subtracted each participant's normed and age-adjusted cognitive composite score from their CFQ score. In contrast, global metacognitive sensitivity was measured by running Pearson's correlations between the CFQ and cognitive composite measures.

[1] Some objective tasks in Mahncke et al., (2021) were excluded because we either failed to find adequate normative data (Ruff Light trails Test) or had non-normative distributions (EXAMINER antisaccade)

Targeted Cognitive Training and Active Control Interventions

Participants were randomized into either 13 weeks of a cognitive training program (BrainHQ) with 23 exercises selected to improve cognitive functions affected by mTBI or an active control[2] condition where participants played 13 computerized puzzle and board games (see Mahncke et al., 2021). The intervention was framed only as cognitive training and there was no mention of depression or PTSD. In the cognitive training intervention, participants self-administered computerized cognitive training for 1 hour each day, 5 days per week, for 13 weeks. All training exercises targeted the speed and accuracy of neural information processing, required attentional focus to perform correctly, and were accompanied by video game-like rewards when trials were performed correctly. Each exercise adapted on a trial-by-trial basis to an individual's current performance with the goal of ensuring users completed ~80% of trials correctly. Participants were also provided specific feedback on their performance and how they were individually improving over the course of the training as well as their current performance in relation to other users. For more detailed information about the exercises and feedback offered, see the Supplementary Materials). In contrast, the active control condition consisted of randomized Hoyle© Puzzle and Board Games selected to minimize demands on the speed and accuracy of information processing. Participants received feedback on their performance (e.g., in-game scoreboards and achievements) as well as in-game rewards (see Supplementary Materials). For both conditions, coaches reviewed progress data regularly and gave telephone-based coaching which provided positive feedback, motivation, and technical support on a weekly basis.

[2] In the clinical trial, this was considered the active control condition. However, because entertainment games could have an impact on clinical/self-reported symptoms (e.g., by correcting metacognitive bias), we label this condition 'active control' here and throughout the manuscript.

Analyses

We first characterized baseline demographic and clinical characteristics by comparing recruitment groups (groups: objective only, self-reported only, and both objective and self-reported inclusion criteria) by running between group ANOVAs for continuous variables and chi-squared tests for categorical variables. We then ran one-sample *t*-tests to determine if norm-adjusted metacognitive bias scores were less than zero (i.e., metacognitive bias in healthy populations). We further examined metacognitive bias differences by running between group ANOVAs for recruitment groups as well as by probable PTSD/depression diagnoses (groups: neither PTSD/depression, depression only, PTSD only, and both PTSD/depression). We additionally ran multiple regressions to determine if depressive and/or PTSD symptoms separately or together explain unique variance in metacognitive bias.

Longitudinal analyses on the effects of cognitive training and the active-control condition were performed in reduced samples because not all participants returned for Time 2 (n=63) or Time 3 (n=45). In calculating a longitudinal difference score in global metacognitive bias and clinical measures of interest, Time 2 scores were subtracted from Time 1 scores, with this method being repeated for Time 3 scores. Changes in global metacognitive bias and self-reported and objective cognition were then associated using Pearson's correlations with changes in clinical measures to see if they tracked with symptom changes and to further explore specificity. Change correlation analyses were repeated with residuals (regressing Time 1 measures out of Time 2) to account for regression to the mean. In order to replicate findings in Agnoli et al., (2023), a joint regression model of changes in depressive/PTSD symptoms predicting changes in metacognitive bias was run to examine if both clinical symptoms explained unique variance in changes in metacognition. Finally, to explore the intervention effects on metacognitive bias,

participants were separated into experimental groups and paired sample *t*-tests for variables of interest and group x pre/post repeated measures ANOVAs were performed to determine if there were main effects or interactions.

It should be noted that we FDR-corrected for multiple comparisons at α=.05 to control for type I errors. Time 1 analyses with a priori hypotheses from Agnoli et al. (2023) were not FDR corrected, except in exploratory ANOVA/chi-squared tests to characterize demographics. All Time 2/Time 3 *t*-tests, ANOVAs, and Pearson correlations were FDR corrected within their set of analyses (e.g., Table 3 change score correlation analyses were FDR-corrected across all 18 correlations, with this method being repeated for the change residuals).

Transparency and Openness

Raw data files are available following standard data sharing protocols at the Boston VA and Posit Science. Computer syntax is available through SPSS and R upon request. The study design and analysis plan were not preregistered.

Sample Size Justification

The current study is part of a previous clinical trial (BRAVE) that found changes in self-reported and objective cognition after an intervention (Mahncke et al., 2021). Our previous study found significant associations between global metacognitive bias and depressive symptoms (N=467, ρ=*-.32,* Agnoli et al., 2023). Because we sought to replicate these previous results using a more sensitive self-report measure of cognition (CFQ), we conservatively expected a similar Pearson's correlation coefficient. Thus, with α=0.05 and $1-\beta$ =0.80, it should only require 71 participants to adequately detect these associations. Therefore, with a sample size of 84 participants, we estimated that we would have enough power to detect an association between metacognitive bias and depressive symptoms.

Results

Demographics, Clinical, and Cognitive Characteristics

Participants (*n*=84, 77% Veterans) were recruited based on self-reported and objective cognitive functioning during their screening visit. Our sample had 57 participants with self-reported cognitive dysfunction and 68 participants with objective cognitive dysfunction, with 41 of those participants meeting recruitment criteria for both self-reported and objective cognitive dysfunction (see Table 1). Notably, participants that met only objective inclusion criteria displayed significantly lower depressive and PTSD symptoms and reduced self-reported cognitive impairments compared to either the self-reported or the self-reported + objective dysfunction group (*p*'s<.003). Additionally, the self-reported + objective dysfunction group was associated with worse objective cognition compared to both groups alone (self-reported inclusion, *t*=2.79, *p*=.007; objective inclusion, *t*=3.15, *p*=.002).

Table 1.

Baseline Demographics and Characteristics

Measure	Full Sample	Self-Reported Inclusion (*n*=16)	Objective Inclusion (*n*=27)	Self-reported and Objective inclusion (*n*=41)			
	Mean (SD)/ Percentages				*F/χ2*	*p*	*q*
Gender (M:F)	69:15	16:0	21:6	32:9	4.30	0.117	0.137
Age	33.61 (8.74)	31.19 (6.94)	31.33 (9.50)	36.05 (8.37)	3.30	0.042	0.053
White	79%	69%	85%	78%			
Black	17%	31%	7%	17%	5.00	0.288	0.288
Other	5%	0%	8%	5%			
Education	14.31 (1.94)	13.75 (1.77)	14.11 (1.78)	14.66 (2.08)	1.48	0.234	0.252

PTSD symptom severity (PCL-C)	45.12 (16.07)	51.56 (17.05)	34.19 (10.98)	49.80 (15.12)	11.67	<0.001	<0.001
Depression symptom severity (BDI-II)	18.83 (12.22)	22.31 (13.52)	11.70 (8.66)	22.17 (11.91)	7.90	<0.001	<0.001
Self-reported cognitive functioning (CFQ; raw)	56.58 (17.92)	59.56 (18.27)	43.74 (13.14)	63.88 (16.12)	13.80	<0.001	<0.001
Self-reported cognitive functioning (CFQ; normed)	-1.84 (1.32)	-1.96 (1.18)	-0.89 (0.99)	-2.42 (1.22)	14.62	<0.001	<0.001
Objective cognitive functioning (battery of cognitive tests; normed)	-0.21 (0.70)	0.06 (.62)	0.04 (0.66)	-0.47 (0.66)	6.73	0.002	0.003
RNBI attention	66.98 (14.14)	70.5 (13.27)	53.44 (9.18)	74.51 (11.12)	30.96	<0.001	<0.001
RNBI executive functioning	59.87 (13.52)	62.31 (13.36)	49.11 (10.57)	66 (10.99)	18.52	<0.001	<0.001
RNBI learning	73.70 (17.08)	78 (15.95)	56.78 (8.62)	83.17 (12.99)	37.94	<0.001	<0.001
RNBI memory	67.05 (17.54)	74.81 (21.57)	50.33 (6.43)	75.02 (12.61)	31.24	<0.001	<0.001
ANAM cognitive composite score	-2.23 (1.73)	-0.09 (0.77)	-2.36 (1.31)	-2.99 (1.57)	25.97	<0.001	<0.001

Note. F/χ2, and *p* values were obtained by a between groups ANOVA/chi-squared tests examining the continuous/categorical variable across the self-reported, objective, and self-reported and objective inclusion groups. Self-reported cognitive functioning was reported as a raw score before age-adjusted norming.

Baseline Global Metacognitive Sensitivity and Bias

Metacognitive Sensitivity Across the Sample

Before examining metacognitive bias, our main variable of interest, we wanted to assess whether self-reported cognitive abilities were associated with objective cognitive performance, i.e., metacognitive sensitivity. When correlating age-normed self-report (CFQ) and objective

cognitive composite scores across the sample at baseline, we found a significant association (r=.25, p=.023), suggesting modest global metacognitive sensitivity similar to other studies in both local (e.g., number/word recall, Hildenbrand and Sanchez, 2022); and global domains (e.g., global health and cognition, Agnoli et al., 2023; Zell and Krizan 2014).

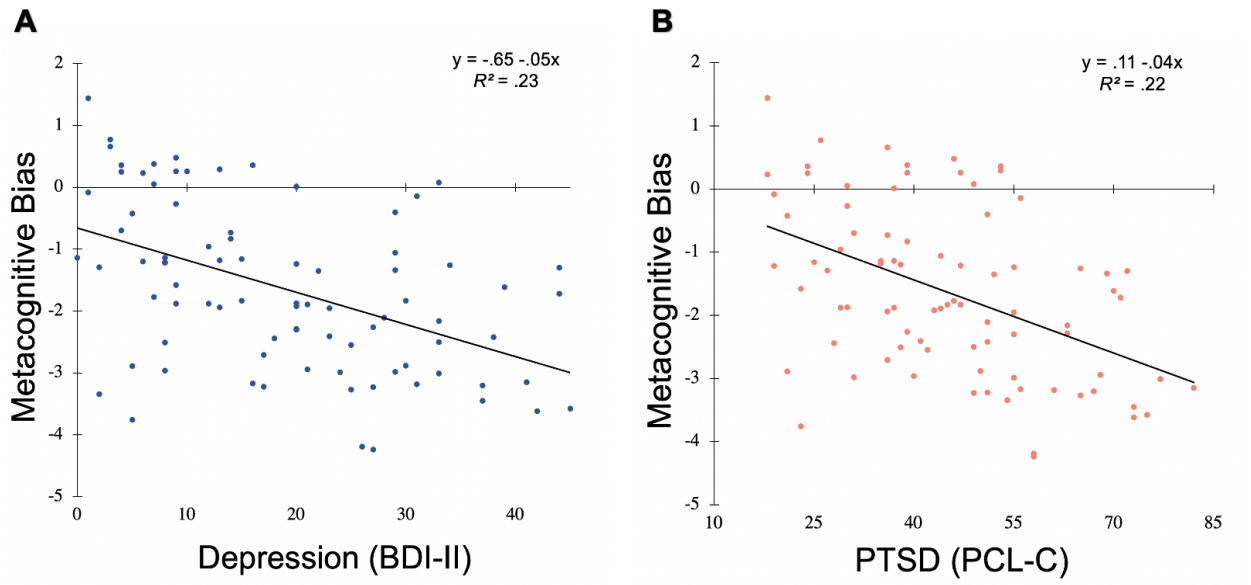
Global Metacognitive Bias

We then examined global metacognitive bias in cognition for each participant, i.e., age-adjusted and normed self-reported - objective performance, and its relationship to depression and PTSD symptoms. The mean of the z-score difference for metacognitive bias in the entire sample was significantly less than zero (M=-1.63, t=-11.26, p<.001), which was driven by large self-reported dysfunction (M=-1.84, SD=1.32, t=-12.79, p<.001) compared to very mild objective cognitive dysfunction (M=-.21, SD=.70 t=-2.74, p=.008). More negative metacognitive biases were significantly associated with depression (r=-.48, p<.001) and PTSD (r=-.47, p<.001; see Figure 2A/B). In a multiple regression, neither depressive symptoms nor PTSD symptoms uniquely predicted metacognitive bias (depression β=-.29, p=.066; PTSD β=-.24, p=.133)[3].

[3] We additionally ran a regression model for depressive/PTSD symptoms predicting self-reported cognition. PTSD symptoms uniquely predicted self-reported cognition (β=-.43, p=.003), though depressive symptoms did not (β=-.22, p=.129)

Figure 2.

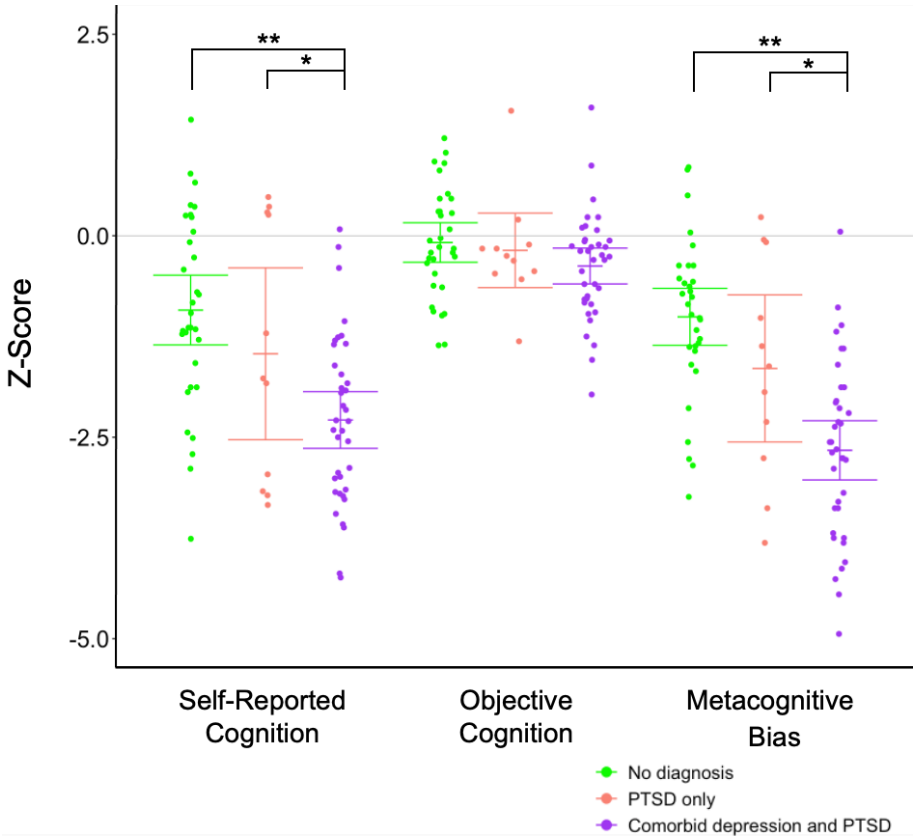
Scatterplot of Metacognitive Bias by A) Depressive and B) PTSD Symptoms



This was consistent when we examined metacognitive biases by probable diagnostic groups (neither PTSD or depression, PTSD or depression only, both PTSD and depression; see Figure 3) and observed that participants with probable depression and PTSD diagnoses had numerically more negative metacognitive biases when compared to those without a probable diagnosis (no diagnosis n=32, *M bias*=-.92, *SD*=1.20; comorbid depression and PTSD n=37, *M bias*=-2.29, *SD*=1.05, t=5.03, p<.001; see Figure 3). Additionally, those with probable PTSD only (*M bias*=-1.47, *SD*=1.59) had less negative metacognitive bias than those with both PTSD and depression (t=2.01, p=.050). The probable depression group (n=4) had a comparably negative metacognitive bias (*M bias*=-1.77, *SD*=1.27).

Figure 3.

Differences in Self-Reported, Objective, and Metacognitive Bias by Probable PTSD/Depression Diagnoses



Note. Diagnoses refer to probable diagnoses based on cut-off scores (for depression, BDI-II scores of ≥19; for PTSD, PCL-C scores of ≥40, see Methods). We removed the depression only group because of its low sample size (n=4) and report it here: self-reported cognition (M=-1.49, SD=1.35), objective cognition (M=.28, SD=.91), and metacognitive bias (M=-1.77, SD=1.27). For the sake of graphical interpretation, self-reported cognition was reverse scored, such that positive values represent improved cognitive functioning relative to normative populations. Error bars represent standard error. **p<.01. *p<.05.

We further characterized metacognitive bias differences across cognitive inclusion criteria (objective dysfunction, self-reported dysfunction, and objective and self-reported dysfunction). Participants with only self-reported dysfunction (*M bias*=-2.02, *SD*=1.02) or self-reported and objective dysfunction (*M bias*=-1.95, *SD*=1.35) had more negative metacognitive biases than those with only objective dysfunction (*M bias*=-.93, *SD*=1.23; *t*'s=2.99, 3.15; *p*'s=.005, .002, respectively).

Effects of Cognitive Training on Metacognitive Bias and Clinical Symptoms

We next sought to address the question of whether cognitive training with extensive feedback can improve metacognitive sensitivity and biases and whether this relates to improvements in depressive/PTSD symptoms. A subsample of participants (*n*=63)[4] completed 5 sessions/week x 13 weeks of either cognitive training or an active control condition (see Figure 1). We began by examining pre/post differences between the experimental training and active control conditions in metacognitive sensitivity. Metacognitive sensitivity generally did not change from Time 1 (*r*=.26, *p*=.042) to Time 2 (*r*=.24, *p*=.060), though it did numerically increase in the further reduced sample (*n*=45) three months after the study ended (Time 3, *r*=.46, *p*=.001).

We next examined pre/post differences in self-reported cognition, objective cognition, and metacognitive bias, FDR-correcting for multiple comparisons[5]. Using the objective cognitive composite (see Methods), we ran a group x pre/post repeated-measures ANOVA and found that the experimental training group showed a trend towards greater improvements than the active

[4] For those who did not complete 13 weeks of cognitive training (*n*=21), their depression, PTSD, and metacognitive bias did not significantly differ from those who did complete treatment at Time 1 (all *p*'s $\geq$ .508).

[5] We additionally measured if baseline clinical or cognitive symptoms significantly predicted changes in depression, PTSD, or metacognitive bias using either change scores or residuals (see Methods), and only Time 1 depressive symptoms marginally predicted changes in depressive symptoms (*r*=.22, *p*=.075).

control group (F=2.42, one-tailed p=.062, similar to Mahncke et al., 2021[6]). However, there were no significant pre/post x group interactions for either self-reported cognitive functioning (F=.15, p=.703) or metacognitive bias (F=.92, p=.222), nor for depression or PTSD (F=.01, 1.08; p=.924, .303, respectively). Notably, all measures except objective cognition showed significant main effects of pre/post training.

When collapsing across training groups, comparing pre vs. post training and FDR correcting for multiple comparison showed that participants' metacognitive bias became significantly more positive (M difference=.56, t=4.01, p<.001, q<.001) and their self-reported cognition improved (M difference=.66, t=5.59, p<.001, q<.001), but there was no significant overall objective cognition improvements (M difference=.10, t=1.23, p=.222). Additionally, participants significantly reduced their overall depressive symptoms (M difference=3.34, t=3.49, p<.001, q=.002, clinically significant effect is $\geq$ 3 points, Button et. al., 2015) and PTSD symptoms (M difference=3.45, t=3.06, p=.003, q=.004, clinically significant effect is $\geq$ 5 points, Khoo et al., 2011). These patterns were present for each group (Cohen's d's$\geq$.37, see Table 2), except for PTSD symptoms in the active control group (Cohen's d=.25, t=1.38, p=.179).

We additionally examined treatment effects in a subset of participants who returned 3 months after the intervention was completed (experimental training n=25; active control n=21). After FDR-correcting for multiple comparisons, both groups continued to show reductions in depressive symptoms (Time 1 vs. Time 3 t's$\geq$2.52, Cohen's d's$\geq$.52, p's$\leq$.019, q's$\leq$.027; see Supplementary Table S1). Consistent with Time 1 vs. Time 2 intervention effects, only the active control group retained their reduction in PTSD (Time 1 vs. Time 3 t=2.90, Cohen's d=.63,

[6] The current study used a subset of the objective cognitive tasks that were normally distributed and had adequate normative data, see methods.

p=.009, *q*=.015). The active control group additionally retained significant improvements in metacognitive bias (Time 1 vs. Time 3 *t*=3.62, Cohen's *d*=.79, *p*=.002, *p*=.005).

Table 2.

Cognitive Training and Active Control Intervention Effects

Measure	Experimental Training (*n*=30)						Active Control (*n*=33)					
	Time 1	Time 2	Cohen's d	*t*	*p*	*q*	Time 1	Time 2	Cohen's d	*t*	*p*	*q*
Depression symptom severity (BDI-II)	16.73 (11.58)	13.47 (10.65)	0.55	3.01	0.005	0.013	21.09 (13.19)	17.68 (14.81)	0.38	2.23	0.032	0.051
PTSD symptom severity (PCL-C)	41.20 (14.00)	39.00 (16.65)	0.25	1.38	0.179	0.199	47.15 (18.21)	42.59 (17.86)	0.52	3.02	0.005	0.014
Metacognitive Bias	-1.17 (1.33)	-0.79 (1.78)	0.37	-2.00	0.055	0.069	-1.99 (1.29)	-1.27 (1.20)	0.63	-3.68	<0.001	0.004
Objective Cognition	-0.33 (0.71)	-0.09 (0.64)	0.40	-2.21	0.035	0.051	-0.17 (0.71)	-0.20 (0.74)	0.05	-0.19	0.876	0.876
Self-reported cognitive functioning (CFQ)	-1.50 (1.40)	-0.89 (1.70)	0.79	-4.34	<0.001	<0.001	-2.16 (1.25)	-1.48 (1.29)	0.65	-3.76	<0.001	0.003

Note. Participants were assigned randomly to the experimental training or active control condition (i.e., not based on recruitment criteria). In particular, 8/13 participants received the intervention with only self-reported cognitive deficits, compared to 9/19 participants with only objective cognitive deficits, and 13/31 with objective and self-reported cognitive deficits.

Associations between Changes in Metacognitive Bias and Changes in Clinical Symptoms

To determine if changes in depressive and PTSD symptoms were associated with changes in metacognitive bias and/or self-reported and objective cognition, we ran Pearson correlations across the full sample and each treatment condition. Across conditions, changes in metacognitive bias from Time 1 to Time 2 were significantly associated with changes in both depressive (*r*=-.41, *p*<.001, *q*<.001) and PTSD symptoms (*r*=-.42, *p*<.001, *q*<.001; see Table 3), such that more positive changes in metacognitive bias were associated with decreases in PTSD and depressive

symptoms. In a multiple regression predicting metacognitive bias, changes in PTSD symptoms explained unique variance in metacognitive bias and changes in depressive symptoms trended towards predicting unique variance (adjusted R^2=.22; PTSD β=-.29, p=.043; depression β=-.24, p=.082).

Table 3.

Cognitive Training and Active Control Intervention Change Correlations

Change Score	Full Sample (n=63)		Intervention (n=30)		Active Control (n=33)	
	Depression	PTSD	Depression	PTSD	Depression	PTSD
Metacognitive bias	-0.41**	-0.42**	-0.18	-0.35	-0.54**	-0.46*
Objective cognition	0.06	0.04	-0.05	0.13	0.11	-0.08
Self-reported cognition	-0.44**	-0.48**	-0.27	-0.37	-0.51**	-0.55**
Residuals	Depression	PTSD	Depression	PTSD	Depression	PTSD
Metacognitive bias	-0.47**	-0.42**	-0.37	-0.35	-0.55**	-0.49**
Objective cognition	0.12	0.04	0.12	0.13	0.12	0.03
Self-reported cognition	-0.48**	-0.47**	-0.40	-0.37	-0.51**	-0.55**

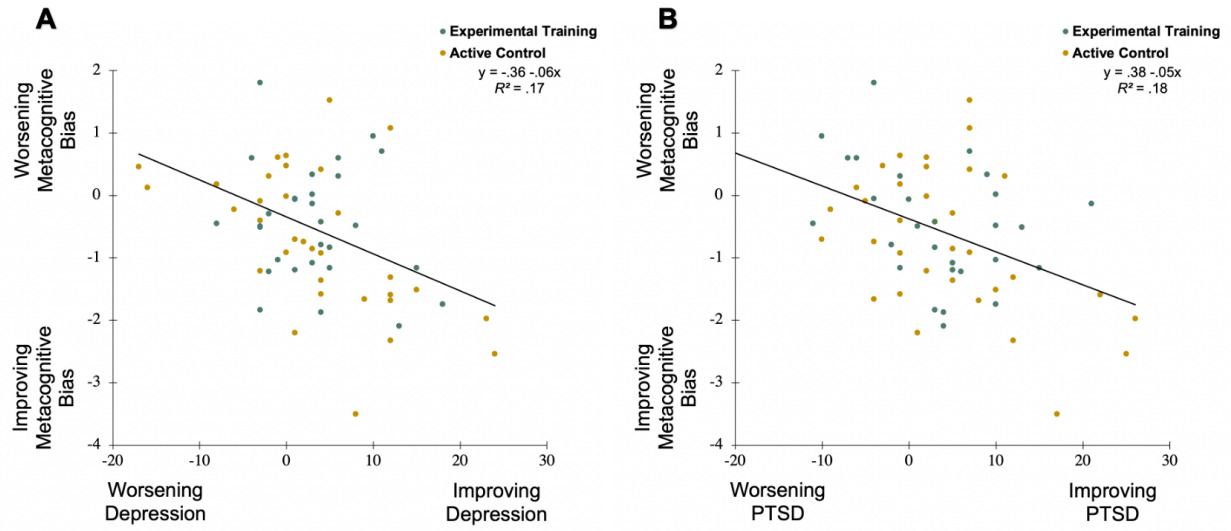
Note. Residuals were calculated by residualizing Time 2 scores after removing the influence of Time 1 in order to control for regression to the mean. Analyses were FDR-corrected across the 18 Pearson's correlations of change scores and again for the residuals. *q=.05. **q=.01.

When examining the training programs separately, the active control group showed significant associations between metacognitive bias changes and reductions in both depressive symptoms (r=-.54, p=.001, q=.003 see Figure 4) and PTSD symptoms (r=-.46, p=.007, q=.016) where in the experimental training these associations were only numerical or trending (depression r=-.18, p=.354, q=.520; PTSD r=-.30, p=.057, q=.103). Additionally, we repeated these correlations using residuals of depression and PTSD symptom changes (Time 2 scores

regressing out Time 1) in order to control for regression to the mean, and in all cases, the observed patterns remained the same or increased in strength (see Table 3).

Figure 4.

Scatterplot of Changes in Metacognitive Bias by Changes in A) Depressive and B) PTSD Symptoms



We additionally correlated changes in both objective and self-reported cognition with changes in depression and PTSD. Notably, changes in objective cognition were not correlated with changes in clinical symptoms in either the difference scores or residuals (all p's>.305) while improvements in self-reported cognition were generally related to reductions in clinical symptoms (e.g., full sample depression r=-.44, p<.001, q=.003; PTSD r=-.48, p<.001, q=.001; see Table 3).

Do Changes in Metacognition and Clinical Symptoms differ by enrollment criteria?

One possibility is that improvements in metacognitive bias and depressive/PTSD symptoms were specific to individuals with subjective cognitive complaints at enrollment. Importantly, all enrollment groups showed more numerical improvements in metacognitive bias

after training (Cohen's *d's*≥.16) and observed numerical reductions in depressive symptoms (all Cohen's *d's*≥.39) and PTSD symptoms (Cohen's *d's*≥.34; see Supplementary Table S2), suggesting this was a more general effect and not specific to groups with subjective cognitive deficits.

Discussion

The present study provides important insights into the relationships between global metacognitive bias and depression/PTSD. First, we observed significantly negative metacognitive biases in a sample with mTBI and high rates of depressive/PTSD symptoms in relation to normative populations, where more negative biases were associated with greater baseline depressive/PTSD symptoms (*r*=-.48/-.47, respectively). The present study also demonstrated that 13 weeks of either targeted or active control (entertainment games) training reduced global negative metacognitive biases, and further, that this reduction was associated with decreases in depressive and PTSD symptoms after training (*r*=-.41/-.42, respectively). Together, these findings have important theoretical implications for the role of global metacognition in depression and PTSD and suggest that feedback-focused cognitive training is a potential method to improve metacognitive biases as well as reduce symptoms of PTSD and depression.

The present results observed a robust relationship between global metacognitive bias and depression/PTSD. We found moderate correlations between metacognitive bias and depressive (*r*=-.48) as well as PTSD symptoms (*r*=-.47). These associations are similar to our previous study with veterans (*ρ*=-.32 and -.23, respectively, Agnoli et al., 2023). Notably, our self-reported and objective cognitive measures differed between studies, suggesting that the global metacognitive bias association with depressive and PTSD symptoms is robust and replicable rather than being dependent on the specific cognitive tasks/questionnaires used. One reason why

these associations were slightly stronger in the present study is likely because we used the 25-item CFQ, which is more reliable and has less of a floor effect than the 6-item WHODAS-II subscale in Agnoli et al. (2023). One interesting difference is that Agnoli et al. (2023) found that depressive symptoms mediated the relationship between metacognitive bias and PTSD symptoms, whereas in the present study PTSD and depression had similar, robust relationships with metacognitive bias, with neither explaining unique variance. With regards to changes in metacognitive biases and changes in depression/PTSD, despite different time scales (13 weeks vs. 2 years, respectively) and that the present study involved interventions, both the present study and Agnoli et al. (2023) found that changes in metacognitive bias were significantly associated with changes in depressive (Previous: ρ=-.25, Current: r=-.41) and PTSD symptom (Previous: ρ=-.33, Current: r=-.42). This provides strong converging evidence that changes in global metacognitive biases are closely linked with changes in depressive and PTSD symptoms.

Beyond reaffirming the strong link between global metacognitive bias and depression/PTSD, the present study is the first to quantify global metacognitive biases in a clinical sample. In our mTBI sample of mostly (77%) veterans with high rates of PTSD and depressive symptoms, we found overall substantial negative metacognitive biases, with the mean z-score difference (M=-1.63) compared to age-adjusted normative populations. In particular, participants self-reported cognitive deficits of 1.84 SDs below normative populations, while the magnitude of their objective cognitive deficits were ~9 x less severe (M *z-score*=-.21). Additionally, the present study observed that this bias was more negative in individuals with probable depression only (M=-1.77), PTSD only (M=-1.47), or depression + PTSD (M=-2.29), than those without either disorder (M=-.92). While depression has been related to negative metacognitive biases in local domains compared to healthy controls (e.g., knowledge

judgements, Fu et al., 2005), it has remained unclear whether the relationship between *global* metacognitive biases and depression and PTSD are negative compared to the general population. Past studies examining metacognitive bias in clinical populations have less precisely measured global metacognition, for example, by comparing general self-esteem on the Rosenberg Self-Esteem Scale (i.e., rather than cognitive self-esteem) to local metacognitive confidence on a cognitive task (Hoven et al., 2022), or by measuring self-reported cognition using a cognitive disability scale which did not have adequate normative data (Agnoli et al., 2023). The present study improves upon previous research by using the CFQ, which globally indexes cognitive abilities and has an adequate range in normative populations (Goodman et al,. 2022). We also improved upon previous work by contrasting the age-normed CFQ with a battery of age-normed cognitive tasks, allowing us to calculate global metacognitive biases in more interpretable z-score units.

In addition to quantifying the degree of metacognitive bias in cognition and its relation to depression/PTSD, to our knowledge the present study is the first to observe improvements in global metacognitive bias from cognitive training interventions. Participants completed 65 one-hour sessions over 13 weeks of either targeted cognitive games which were designed to improve perception, attention, and memory and provided adaptive feedback (see Supplementary Table S1), or entertainment games with minimal cognitive demands (e.g., Tetris-clone, Mahjong, and Solitaire). Notably, while objective cognition numerically improved only in the experimental training condition, both the experimental training and active control groups improved in self-reported cognition and improved their negative metacognitive biases with a medium effect size (Cohen's *d*=.37 and .64, respectively) in addition to reductions in depressive/PTSD symptoms, with these benefits generally persisting 3 months later. This is consistent with other targeted and

entertainment cognitive training studies that, despite not explicitly measuring metacognitive biases, showed greater improvements in subjective than objective cognitive measures (e.g., Cohen-Mansfield et al., 2014; Baniqued et al., 2014).

Importantly, in the current study we found that the degree to which individuals improved their metacognitive bias was related to improvements in depressive and PTSD symptoms across both training conditions, with numerically stronger effects after entertainment game training. These results align with a recent review finding benefits of entertainment games on mood and anxiety symptoms (Boldi and Rapp, 2022). One plausible interpretation of the clinical benefits we observed is that the extensive feedback provided may have recalibrated participants' metacognitive bias, e.g., by providing more salient memories of successfully accomplishing cognitive tasks (Dalgleish and Werner-Seidler, 2014). These salient successes may overcome the tendency to retrieve negative self-referential memories in depression (Marchetti et al., 2018; Duyser et al., 2020), leading to reduced metacognitive bias and depressive symptom improvements by ameliorating negative schemas (Rude et al., 2003). Though future studies are necessary to provide support for this interpretation, the fact that effect sizes of training were greater for improvements in metacognitive bias than for depression/PTSD preliminarily supports that reducing negative metacognitive biases may have been a more primary effect of training. An alternative explanation is that training may have directly improved depressive and PTSD symptoms by providing a distraction (Hemenover et al., 2018) or by improving general self-efficacy (not specific to training/games) that reduced both depressive/PTSD symptoms and metacognitive bias. A future study including cognitive training without any feedback vs. intensive feedback (providing greater recalibration of metacognitive bias) along with a test-retest

control would be useful to test the relationship of metacognitive bias improvements with depressive and PTSD symptom improvements.

Though it would be important to replicate the current results and further test the mechanisms of action, they do provide preliminary evidence for interventions that target the modification of metacognitive biases to improve clinical outcomes. This approach could be adapted to additional domains (e.g., physical training/exercise, social cognitive abilities), which may have even larger impacts on depressive and PTSD symptoms compared to cognitive training alone. This approach is in accordance with recent clinical trials examining metacognitive training for depression (D-MCT), a variation of CBT which seeks to improve metacognitive sensitivity and bias on various local cognitive tasks and abilities, as well as dysfunctional beliefs about one's thought processes (e.g., rumination, thought suppression, social withdrawal) in depression. D-MCT has observed reductions in false memories (Moritz et al., 2018), belief inflexibility (So et al., 2021), and depressive symptoms generally (Jelinek et al., 2019; Jelinek et al., 2016). One study sought to delineate the mechanisms of change between depression and metacognitive awareness by separating modules of D-MCT. They found that, compared to other D-MCT modules, metacognitive awareness training focused on self-esteem, memory deficits, and confidence in emotion recognition showed greater improvements in depressive symptoms (Miegel et al., 2023). Incorporating these insights from the current and past studies, future interventions in depression and PTSD might seek to optimize objective cognitive feedback in order to correct global negative metacognitive biases. For example, targeted games could balance adaptive difficulty with gamified motivation in order to provide individuals with a feeling of accomplishment. Additionally, these games might provide transparent performance measurements, and/or compare patients' abilities to normative distributions. These methods of

feedback were provided in the current study's experimental training and, encouragingly, are widely employed in commercial games, where players unlock more difficult levels, are provided comprehensive performance statistics, and can be normatively ranked, which rankings are associated with objective cognitive ability (Kokkinakis et al. 2017). Future therapies targeting metacognition may be particularly beneficial in samples with large negative metacognitive bias, such as veterans with mTBI.

Though the current findings replicate previous results and provide several novel insights, there are limitations. Participants were recruited due to their self-reported or objective cognitive impairment and a history of mTBI. This sample may be particularly prone to negative metacognitive biases (e.g., 'I haven't been the same since my traumatic brain injury') and the results may not generalize to other populations with fewer subjective complaints. Additional generalization concerns include that participants were primarily veterans who were predominantly white and male, though our sample included ~23% civilians and had large ranges in age (18-50). The sample also had considerable comorbidity between PTSD and depression, whereas samples with greater clinical separation would be useful to dissociate the effects of PTSD from depression. Though the current study is the first to measure global metacognitive bias compared to normative populations, future studies may benefit from more comprehensively measuring metacognition at both local domains (e.g., obtaining trial-by-trial confidence ratings within each cognitive task) and global domains. Another issue is that in finding effects in both the experimental and active control cognitive training groups, we lacked a proper control group, making it challenging to identify the causal mechanisms of improvements in both metacognitive bias and depressive/PTSD symptoms. Finally, because metacognitive bias and subjective cognitive complaints were highly correlated in the current sample, we were not able to

characterize whether they have differential associations/effects of cognitive training, which would be an important endeavor for future studies.

Summary

The current study is the first to quantify the magnitude of global metacognitive bias by contrasting self-reported and objective cognitive measures in relation to normative populations. We replicated previous research finding that global negative metacognitive biases are robustly related to PTSD and depressive symptoms. Additionally, we found that global metacognitive biases and depression/PTSD improved after both targeted and entertainment cognitive training. Further, positive changes in metacognitive bias were associated with improvements in PTSD/depressive symptoms. The current findings provide additional evidence for the role of metacognitive bias in depression and PTSD as well as outlining the potential benefits of novel interventions targeting global negative metacognitive biases in populations with cognitive complaints and/or depression, PTSD, and mTBI.

Funding

This work was supported by the US Department of Veterans Affairs through the Translational Research Center for TBI and Stress Disorders, a VA Rehabilitation Research & Development TBI National Research Center (B3001-C). HM, SJ, and the BRAVE study were funded by a CDMRP grant (PT100024) made to Posit Science as the coordinating center, with subcontracts to each of the individual trial sites.

Author Contributions

JD and SA conceived of the study. SA performed the data analysis and visualization under the supervision of JD. SA and JD drafted the paper, and ME, HM, SG, and WM provided

critical revisions. WM contributed funding. All authors approved of the final version of the paper for submission.

Conflicts of Interest

The authors report no conflicts of interest.

References

Abd-Alrazaq, A., Al-Jafar, E., Alajlani, M., Toro, C., Alhuwail, D., Ahmed, A., ... & Househ, M. (2022). The effectiveness of serious games for alleviating depression: systematic review and meta-analysis. *JMIR Serious Games*, *10*(1), e32331.

Achtman, R. L., Green, C. S., & Bavelier, D. (2008). Video games as a tool to train visual skills. *Restorative neurology and neuroscience*, *26*(4-5), 435-446.

Agnoli, S., Zuberer, A., Nanni-Zepeda, M., McGlinchey, R. E., Milberg, W. P., Esterman, M., & DeGutis, J. (2023). Depressive Symptoms are Associated with More Negative Global Metacognitive Biases in Combat Veterans, and Biases Covary with Symptom Changes over Time. *Depression and Anxiety*, *2023*.

Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. *The self in social judgment*, *1*, 85-106.

Baniqued, P. L., Kranz, M. B., Voss, M. W., Lee, H., Cosman, J. D., Severson, J., & Kramer, A. F. (2014). Cognitive training with casual video games: points to consider. *Frontiers in psychology*, *4*, 1010.

Beck, A. T. (1970). Cognitive therapy: Nature and relation to behavior therapy. *Behavior therapy*, *1*(2), 184-200.

Benwell, C. S., Mohr, G., Wallberg, J., Kouadio, A., & Ince, R. A. (2022). Psychiatrically relevant signatures of domain-general decision-making and metacognition in the general population. *npj Mental Health Research*, *1*(1), 10.

Boldi, A., & Rapp, A. (2022). Commercial video games as a resource for mental health: A systematic literature review. *Behaviour & Information Technology*, *41*(12), 2654-2690.

Calkins, A. W., McMorran, K. E., Siegle, G. J., & Otto, M. W. (2015). The effects of computerized cognitive control training on community adults with depressed mood. *Behavioural and cognitive psychotherapy*, *43*(5), 578-589.

Capobianco, L., Heal, C., Bright, M., & Wells, A. (2019). What comes first metacognition or negative emotion? A test of temporal precedence. *Frontiers in psychology*, *10*, 2507.

Cohen-Mansfield, J., Cohen, R., Buettner, L., Eyal, N., Jakobovits, H., Rebok, G., ... & Sternberg, S. (2015). Interventions for older persons reporting memory difficulties: a randomized controlled pilot study. *International Journal of Geriatric Psychiatry*, *30*(5), 478-486.

Dalgleish, T., & Werner-Seidler, A. (2014). Disruptions in autobiographical memory processing in depression and the emergence of memory therapeutics. *Trends in cognitive sciences*, *18*(11), 596-604.

Dos Santos, H., Bredehoft, M. D., Gonzalez, F. M., & Montgomery, S. (2016). Exercise video games and exercise self-efficacy in children. *Global pediatric health*, *3*, 2333794X16644139.

Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. Sage Publications.

Duyser, F. A., Van Eijndhoven, P. F. P., Bergman, M. A., Collard, R. M., Schene, A. H., Tendolkar, I., & Vrijsen, J. N. (2020). Negative memory bias as a transdiagnostic cognitive marker for depression symptom severity. *Journal of Affective Disorders*, *274*, 1165-1172.

Elgamal, S., McKinnon, M. C., Ramakrishnan, K., Joffe, R. T., & MacQueen, G. (2007). Successful computer-assisted cognitive remediation therapy in patients with unipolar depression: a proof of principle study. *Psychological medicine*, *37*(9), 1229-1238.

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in human neuroscience*, *8*, 443.

Fleming, T. M., Bavin, L., Stasiak, K., Hermansson-Webb, E., Merry, S. N., Cheek, C., ... & Hetrick, S. (2017). Serious games and gamification for mental health: current status and promising directions. *Frontiers in psychiatry*, *7*, 215.

Fox, C. A., Lee, C. T., Hanlon, A., Seow, T., Lynch, K., Harty, S., ... & Gillan, C. (2023). Metacognition in anxious-depression is state-dependent: an observational treatment study.

Fu, T., Koutstaal, W., Fu, C. H., Poon, L., & Cleare, A. J. (2005). Depression, confidence, and decision: Evidence against depressive realism. *Journal of Psychopathology and Behavioral Assessment*, *27*, 243-252.

Gascoine, L., Higgins, S., & Wall, K. (2017). The assessment of metacognition in children aged 4–16 years: a systematic review. *Review of Education*, *5*(1), 3-57.

Geers, A. L., Weiland, P. E., Kosbab, K., Landry, S. J., & Helfer, S. G. (2005). Goal activation, expectations, and the placebo effect. *Journal of personality and social psychology*, *89*(2), 143.

Gentile, D. A., Anderson, C. A., Yukawa, S., Ihori, N., Saleem, M., Ming, L. K., ... & Sakamoto, A. (2009). The effects of prosocial video games on prosocial behaviors: International evidence from correlational, longitudinal, and experimental studies. *Personality and Social Psychology Bulletin*, *35*(6), 752-763.

Gunning, F. M., Anguera, J. A., Victoria, L. W., & Areán, P. A. (2021). A digital intervention targeting cognitive control network dysfunction in middle age and older adults with major depression. *Translational psychiatry*, *11*(1), 269.

Granic, I., Lobel, A., & Engels, R. C. (2014). The benefits of playing video games. *American psychologist*, *69*(1), 66.

Hassan, L., Dias, A., & Hamari, J. (2019). How motivational feedback increases user's benefits and continued use: A study on gamification, quantified-self and social networking. *International Journal of Information Management*, *46*, 151-162.

Hemenover, S. H., & Bowman, N. D. (2018). Video games, emotion, and emotion regulation: Expanding the scope. *Annals of the International Communication Association*, *42*(2), 125-143.

Hildenbrand, L., & Sanchez, C. A. (2022). Metacognitive accuracy across cognitive and physical task domains. *Psychonomic Bulletin & Review*, *29*(4), 1524-1530.

Hoge, C. W., Auchterlonie, J. L., & Milliken, C. S. (2006). Mental health problems, use of mental health services, and attrition from military service after returning from deployment to Iraq or Afghanistan. *Jama*, *295*(9), 1023-1032.

Hoven, M., Luigjes, J., Denys, D., Rouault, M., & van Holst, R. J. (2023). How do confidence and self-beliefs relate in psychopathology: A transdiagnostic approach. *Nature Mental Health*, 1-9.

Karstoft, K. I., Andersen, S. B., Bertelsen, M., & Madsen, T. (2014). Diagnostic accuracy of the Posttraumatic Stress Disorder Checklist–Civilian Version in a representative military sample. *Psychological Assessment*, *26*(1), 321.

Khoo, A., Dent, M. T., & Oei, T. P. (2011). Group cognitive behaviour therapy for military service-related post-traumatic stress disorder: Effectiveness, sustainability and repeatability. *Australian & New Zealand Journal of Psychiatry*, *45*(8), 663-672.

Kokkinakis, A. V., Cowling, P. I., Drachen, A., & Wade, A. R. (2017). Exploring the relationship between video game expertise and fluid intelligence. *PloS one*, *12*(11), e0186621.

Lyons, E. J. (2015). Cultivating engagement and enjoyment in exergames using feedback, challenge, and rewards. *Games for health journal*, *4*(1), 12-18.

Maniscalco, B., Charles, L., & Peters, M. A. (2022). Optimal Metacognitive Decision Strategies in Signal Detection Theory.

Mahncke, H. W., DeGutis, J., Levin, H., Newsome, M. R., Bell, M. D., Grills, C., ... & Merzenich, M. M. (2021). A randomized clinical trial of plasticity-based cognitive training in mild traumatic brain injury. *Brain*, *144*(7), 1994-2008.

Manos, R. C., Rusch, L. C., Kanter, J. W., & Clifford, L. M. (2009). Depression self-stigma as a mediator of the relationship between depression severity and avoidance. *Journal of Social and Clinical Psychology*, *28*(9), 1128-1143.

Marchetti, I., Everaert, J., Dainer-Best, J., Loeys, T., Beevers, C. G., & Koster, E. H. (2018). Specificity and overlap of attention and memory biases in depression. *Journal of affective disorders*, *225*, 404-412.

Merry, S. N., Stasiak, K., Shepherd, M., Frampton, C., Fleming, T., & Lucassen, M. F. (2012). The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. *Bmj*, *344*.

Miegel, F., Cludius, B., Hottenrott, B., Demiralay, C., Sure, A., & Jelinek, L. (2020). Session-specific effects of the metacognitive training for obsessive-compulsive disorder (MCT-OCD). *Psychotherapy research*, *30*(4), 474-486.

Moroz, M., & Dunkley, D. M. (2015). Self-critical perfectionism and depressive symptoms: Low self-esteem and experiential avoidance as mediators. *Personality and Individual Differences*, *87*, 174-179.

Motter, J. N., Pimontel, M. A., Rindskopf, D., Devanand, D. P., Doraiswamy, P. M., & Sneed, J. R. (2016). Computerized cognitive training and functional recovery in major depressive disorder: a meta-analysis. *Journal of affective disorders*, *189*, 184-191.

Naismith, S. L., Diamond, K., Carter, P. E., Norrie, L. M., Redoblado-Hodge, M. A., Lewis, S. J., & Hickie, I. B. (2011). Enhancing memory in late-life depression: the effects of a combined psychoeducation and cognitive training program. *The American Journal of Geriatric Psychiatry*, *19*(3), 240-248.

Nelson, T. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. Psychol. Bull. 95, 109–133. doi: 10.1037/0033-2909.95.1.109

Orth, U., & Robins, R. W. (2013). Understanding the link between low self-esteem and depression. *Current directions in psychological science*, *22*(6), 455-460.

Pavlas, D., Heyne, K., Bedwell, W., Lazzara, E., & Salas, E. (2010, September). Game-based learning: The impact of flow state and videogame self-efficacy. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 54, No. 28, pp. 2398-2402). Sage CA: Los Angeles, CA: SAGE Publications.

Reeves, D. L., Winter, K. P., Bleiberg, J., & Kane, R. L. (2007). ANAM® Genogram: Historical perspectives, description, and current endeavors. *Archives of Clinical Neuropsychology*, *22*(Suppl_1), S15-S37.

Riley, E., Mitko, A., Stumps, A., Robinson, M., Milberg, W., McGlinchey, R., ... & DeGutis, J. (2019). Clinically significant cognitive dysfunction in OEF/OIF/OND veterans: Prevalence and clinical associations. *Neuropsychology*, *33*(4), 534.

Rouault, M., Seow, T., Gillan, C. M., & Fleming, S. M. (2018). Psychiatric symptom dimensions are associated with dissociable shifts in metacognition but not task performance. *Biological psychiatry*, *84*(6), 443-451.

Rouault, M., Dayan, P., & Fleming, S. M. (2019). Forming global estimates of self-performance from local confidence. *Nature communications*, *10*(1), 1141.

Rouault, M., Will, G. J., Fleming, S. M., & Dolan, R. J. (2022). Low self-esteem and the formation of global self-performance estimates in emerging adulthood. *Translational Psychiatry*, *12*(1), 272.

Rude, S. S., Valdez, C. R., Odom, S., & Ebrahimi, A. (2003). Negative cognitive biases predict subsequent depression. *Cognitive Therapy and Research*, *27*, 415-429.

Sacher, M., Tudorache, A. C., Clarys, D., Boudjarane, M., Landré, L., & El-Hage, W. (2018). Prospective and retrospective episodic metamemory in posttraumatic stress disorder. *Journal of Clinical and Experimental Neuropsychology*, *40*(9), 865-873.

Schmidt, M. (1996). *Rey auditory verbal learning test: RAVLT: a handbook*. Western Psychological Services.

Seal, K. H., Bertenthal, D., Samuelson, K., Maguen, S., Kumar, S., & Vasterling, J. J. (2016). Association between mild traumatic brain injury and mental health problems and self-reported cognitive dysfunction in Iraq and Afghanistan Veterans. *Journal of Rehabilitation Research & Development*, *53*(2).

Seow, T. X., Rouault, M., Gillan, C. M., & Fleming, S. M. (2021). How local and global metacognition shape mental health. *Biological psychiatry*, *90*(7), 436-446.

Soderstrom, N. C., Davalos, D. B., & Vázquez, S. M. (2011). Metacognition and depressive realism: Evidence for the level-of-depression account. *Cognitive neuropsychiatry*, *16*(5), 461-472.

Stricker, N. H., Christianson, T. J., Lundt, E. S., Alden, E. C., Machulda, M. M., Fields, J. A., ... & Petersen, R. C. (2021). Mayo normative studies: regression-based normative data for the auditory verbal learning test for ages 30–91 years and the importance of adjusting for sex. *Journal of the International Neuropsychological Society*, *27*(3), 211-226.

Wechsler, D. (2008). Wechsler adult intelligence scale–Fourth Edition (WAIS–IV). *San Antonio, TX: NCS Pearson*, *22*(498), 1.

Wechsler, D. (1997). *WAIS-3., WMS-3: Wechsler adult intelligence scale, Wechsler memory scale: Technical manual*. Psychological Corporation.

Zell, E., & Krizan, Z. (2014). Do people have insight into their abilities? A metasynthesis. *Perspectives on Psychological Science*, *9*(2), 111-125.

Supplementary Methods

Calculating Age-Adjusted Normative Data

In order to calculate a metacognitive bias z-score, we selected self-reported and objective cognitive measures that 1) were normally distributed, and 2) had adequate normative data (acquirable through either contacting the authors or published) which was further segmented around our population age range of 18–50. Z-scores were then computed within the available age range for each measure (e.g., RAVLT age range of 20–30 years; *M*=52.3, *SD*=7.7, Stricker et al., 2021) before subtracting the normalized objective cognitive composite measure from the self-reported cognitive functioning measure (CFQ). For the CFQ, data was obtained by contacting Goodman and colleagues (2022), who provided means and standard deviations from the Nathan Kline Institute-Rockland Sample (*n*=839) for the following age brackets: 16-19, 20–29, 30–39, and 40–50. Digit Span was normalized using the Wechsler Adult Intelligence Scale Scoring Manual (2008). Symbol Span was normalized using the Wechsler Memory Scale Scoring Manual (WMS-IV; 1997). Both the Flanker and Set-Shifting tasks norms were obtained by contacting Kramer and colleagues for normative distributional data on the Executive Abilities: Measures and Instruments for Neurobehavioral Evaluation and Research (*n*=1515; EXAMINER; 2014), who provided means and standard deviations for the following age brackets: <26, 26–33, 34–42, and >42. Finally, the Rey Auditory Learning Test (RAVLT) sum of trials 1–5 and delayed recall were obtained from a normative study of 4428 cognitively unimpaired adults (Stricker et al., 2021) and were separated into the following age brackets: 16-19, 20–29, 30–39, and $\geq$40.

Cognitive Training and Feedback Mechanisms

The cognitive training exercises provided real-time feedback as the exercises were performed, and delayed feedback in the form of information about progress and performance. Each exercise was composed of a number of individual trials, where a single trial was similar to a trial in a psychophysical task involving stimulus presentation and user response. If the response was correct, users were rewarded with real-time auditory feedback (a "ding" for correct, and a "thonk" for incorrect) and visual animations (e.g., fireworks). After the number of correct trials was performed in the M-up/N-down adaptive algorithm, a challenge meter animated and advanced to show the user that they had advanced to a higher level of difficulty. At the end of a block of trials, quantitative feedback was provided to the user in the form of stars, based on the user's performance relative to a large database of other users who had trained with BrainHQ, with 1 star indicating performance of -0.5 standard deviations below then mean, 2 stars is between -0.5 and +0.5 standard deviations of the mean, 3 stars between +0.5 and +1 standard deviations above the mean, 4 stars between +1 and +1.5 standard deviations above the mean, 5 stars +1.5 standard deviations or more above the mean. For example, after completing the visual perception task 'Hawk Eye,' participants might receive: "★★☆☆☆; your speed was 562 milliseconds, your best is 382 milliseconds; as you improved, the birds appeared for fewer milliseconds. Your score shows how quickly you could take in and accurately respond to what you saw."

Following the completion of all training for the day, users were provided calendars and summaries where they could review their usage (days trained, minutes/hours trained), their progress (the total number of levels they had completed, the levels completed in each exercise), and their overall performance based on percentile scores. Percentile scores were calculated by

averaging their scores on each training level within an exercise relative to the normative database in order to create exercise percentile scores. Exercise percentile scores were then averaged into domain percentile scores (e.g., memory, speed, attention). Percentile scores were calculated for the first time each training level was performed (presented to users as a baseline score) and for the best performance on each training level (allowing a percentile improvement score to be calculated and shown).

Active Control and Feedback Mechanisms

The Active Control training exercises included 13 off-the-shelf computer games like Tetris, Mahjong, Checkers, Go, Solitaire, and Hangman chosen from the 2009 version of HOYLE© Puzzle and Board Games to match the Experimental Training program in the intensity and duration of the training. Active Control training games did not include the adaptive performance features like the Experimental Training exercises, and therefore did not receive detailed performance feedback. In addition, participants did not receive normative performance feedback comparing their performance to others.

However, participants did receive real-time gamified feedback in the form of visible scoreboards which tracked their performance within the context of each game, as well as achievement-like events, which earned them in-game currency called HOYLE bucks. For example, participants would earn points for preventing rows of blocks from stacking up (in Panic, a Tetris clone), removing the opponent's pieces from the board (in Checkers), or figuring out the secret word (in Hangman). The more the participants earned points or achievements, the more they earned HOYLE Bucks. Participants were then able to spend their bucks on new card decks, backgrounds, music tracks, and décor items within the game's catalog. A participant's HOYLE Bucks balance and any achievements they accomplished were shown in the HOYLE

Bucks display in the lower-right corner of every game screen. Rolling over the display showed the four most recent achievements accomplished in that game.

Persistence of Cognitive Training Treatment 3-Months Later

In a subset of participants (*n*=46), we collected data 3-months after cognitive training. Both groups retained their reduction in depression symptoms, as well as improvements in self-reported cognition, though only the active control group retained their reduction in PTSD symptoms and improvements in metacognitive bias (see Table 3).

Table S1.
Cognitive Training Intervention 3-Months Later

Measure	Experimental Training (n = 25)							Active Control (n = 21)						
	Time 1	Time 2	Time 3	Cohen's *d*	*t*	*p*	*q*	Time 1	Time 2	Time 3	Cohen's *d*	*t*	*p*	*q*
Depression symptom severity (BDI-II)	16.21 (11.87)	13.00 (11.24)	12.96 (12.25)	0.52	2.52	0.019	0.027	24.57 (14.10)	21.14 (16.17)	19.90 (16.16)	0.70	3.21	0.004	0.008
PTSD symptom severity (PCL-C)	41.67 (14.62)	40.25 (17.26)	41.83 (18.12)	-0.02	-0.09	0.925	0.925	51.43 (17.71)	44.81 (18.41)	44.00 (19.30)	0.63	2.90	0.009	0.015
Metacognitive Bias	-1.25 (1.42)	-0.88 (1.51)	-1.06 (1.60)	-0.23	-1.14	0.264	0.293	-2.38 (1.26)	-1.61 (1.15)	-1.60 (1.12)	-0.79	-3.62	0.002	0.005
Objective Cognition	-0.32 (0.76)	-0.03 (0.69)	0.12 (0.71)	-0.87	-4.26	<.001	<0.001	-0.14 (0.72)	-0.09 (0.72)	0.04 (0.90)	-0.40	-1.82	0.084	0.105
Self-reported cognitive functioning (CFQ)	-1.57 (1.54)	-0.91 (1.82)	-0.97 (1.72)	-0.93	-4.56	<.001	0.001	-2.52 (1.17)	-1.70 (1.42)	-1.57 (1.37)	-1.00	-4.56	<0.001	0.001

Note. Cohen's *d*, *t*, *p*, and *q* values refer to Time 1 – Time 3 paired samples *t*-tests. The table displays a subsample of participants who returned for a follow-up (Time 3) visit 3-months after the intervention ended.

Recruitment Criteria Effects on Changes in Metacognition and Clinical Symptoms

In exploratory analyses, we examined the relationship between self-reported and objective inclusion criteria on intervention effects in order to determine if baseline recruitment criteria were related to functional outcome. Participants were assigned randomly to the intervention or entertainment training condition (i.e., not based on recruitment criteria)[7]. All groups saw numerical improvements in depressive symptoms of ~3 points, as well as improvements on PTSD (see Table S2). While the only self-reported and only objective inclusion groups saw significant improvements in metacognitive bias (Cohen's *d*'s=1.41, .83; *p's*< .001, .002, respectively), the objective and self-reported inclusion group did not (Cohen's *d*=.16, *p*=.393).

Table S2.

Cognitive Training Intervention by Recruitment Inclusion criteria

Measure	Self-Reported Inclusion (*n*=13)					Objective Inclusion (*n*=19)					Self-Reported and Objective Inclusion (*n*=31)				
	Time 1	Time 2	Cohen's *d*	*t*	*p*	Time 1	Time 2	Cohen's d	*t*	*p*	Time 1	Time 2	Cohen's *d*	*t*	*p*
Depression symptom severity (BDI-II)	22.32 (14.48)	17.85 (13.96)	0.53	1.90	0.082	12.47 (9.50)	9.74 (9.47)	0.47	2.04	0.056	21.71 (12.41)	18.42 (14.01)	0.39	2.17	0.038
PTSD symptom severity (PCL-C)	50.85 (17.80)	45.92 (17.65)	0.45	1.63	0.130	33.47 (11.67)	31.11 (13.20)	0.34	1.49	0.154	48.48 (15.89)	45.06 (17.46)	0.37	2.08	0.046
Metacognitive Bias	-2.06 (1.11)	-0.98 (1.14)	1.41	-5.10	<0.001	-0.88 (1.24)	-0.08 (1.09)	0.83	-3.60	0.002	-1.81 (1.40)	-1.63 (1.60)	0.16	-0.87	0.393
Objective Cognition	0.06 (0.62)	-0.01 (.51)	0.16	-0.58	0.575	-0.05 (0.51)	-0.02 (0.63)	0.06	-0.24	0.814	-0.53 (0.73)	-0.32 (0.77)	0.28	-1.57	0.128
Self-reported cognitive functioning (CFQ)	-2.00 (1.26)	-0.99 (1.39	-1.76	-6.34	<0.001	-0.93 (1.14)	-0.11 (1.16)	0.82	-3.56	0.002	-2.35 (1.28)	-1.94 (1.36)	0.42	-2.36	0.025

[7] In particular, 8/13 veterans received the intervention with only self-reported cognitive deficits, compared to 9/19 veterans with only objective cognitive deficits, and 13/31 with objective and self-reported cognitive deficits.