Attentional bias in depression: understanding mechanisms to improve training and treatment
Anne C Mennen¹, Kenneth A Norman¹,² and Nicholas B Turk-Browne³

One of the most common symptoms of depression is the tendency to attend to negative stimuli in the world and negative thoughts in mind. This symptom is especially nefarious because it is also a cause — biasing processing to negatively valenced information, thus worsening mood, and exacerbating the condition. Here we attempt to systematize the diverse body of recent research on the negative attentional bias from across cognitive and clinical psychology in order to identify recurring themes and devise potential mechanistic explanations. We leverage theoretical progress in our understanding of healthy attention systems in terms of internal versus external components. With this lens, we review approaches to training attention that might reduce the negative attentional bias, including behavioral interventions and real-time neurofeedback. Although extant findings are somewhat mixed, these approaches provide hope and clues for the next generation of treatments.

Addresses
¹Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08540, United States
²Department of Psychology, Princeton University, Princeton, NJ 08540, United States
³Department of Psychology, Yale University, New Haven, CT 06520, United States

Corresponding author: Mennen, Anne C (amennen@princeton.edu)

Current Opinion in Psychology 2019, 29:266–273
This review comes from a themed issue on Attention and perception
Edited by Sarah Shomstein, Andrew Leber and Joy Geng

https://doi.org/10.1016/j.copsyc.2019.07.036
2352-250X © 2019 Elsevier Ltd. All rights reserved.

Introduction
Negative attentional bias in depression — the tendency of depressed individuals to focus on negative stimuli and thoughts more than healthy individuals — has been the subject of research spanning multiple decades, much of it seeking to explain why the bias is not always observed [1–3]. This topic continues to inspire the development of new tasks and training paradigms that have leveraged insights from prior research and new technology. Here we review recent work in this field and explore how basic cognitive investigations of changes in perception and attention in depression may lead to progress in understanding the nature of the bias, and consequently, potential treatment approaches for ameliorating it.

Internal attention
To begin characterizing the negative attentional bias and addressing inconsistencies in the literature, we leverage the framework of internal versus external attention [4]. External attention involves allocating attention to stimuli in the current sensory environment, whereas internal attention operates over mental representations not directly supported by the environment, such as thoughts, emotions, memories, and tasks. Depressed patients perform differently from healthy individuals in both domains. In this section we address different forms of internal attention and related training approaches, before turning to external attention in the next section. First, we highlight two processes that involve attention to internal representations: cognitive control and emotion regulation.

Cognitive control
Relative to healthy individuals, depressed individuals show general deficits in executive functions such as rule shifting, verbal working memory, updating, planning, verbal fluency, psychomotor speed, and especially inhibition [5*,6]. This behavioral difference is often attributed to prefrontal hypoactivity [6] (Figure 1), as depressed individuals perform worse than healthy controls on cognitive control tasks when dorsolateral prefrontal cortex (DLPFC) activity is lower than controls, but not when DLPFC activity is greater than controls [6–8]. Impairments in cognitive control are less robust for at-risk subjects (those identified by a dysphoric mood and high tendency to ruminate) who are completing tasks with neutral materials, but they are consistently found in tasks employing negatively valenced stimuli [5*]. Additionally, deficits in memory performance are most apparent when tasks are relatively unconstrained by instructions, which is thought to allow depressed individuals to spontaneously ruminate on negative thoughts, especially about themselves [1].

These executive deficits associated with depression led to the development of cognitive control training (CCT) — multiple training sessions on tasks devised to improve executive function by demanding inhibition of habitual responses. Tasks that require increased focus on neutral
Neural circuit illustrating how both stimulus-driven and goal-directed pathways can drive the negative attentional bias in depressed subjects (dashed lines indicate weakened functional connections in depression). Negative stimuli trigger a cascade of increased activity (red) from thalamus → amygdala → subgenual cingulate. The DLPFC and thus ACC respond with reduced activity (blue), which in turn reduces inhibition of the amygdala and prolongs processing of negative stimuli. Reprinted with permission from Ref. [52].

stimuli, such as sounds and numbers, reduce rumination [9] and depressive symptoms [9,10], though not always in a generalized manner [56]. A recent review that evaluated CCT across different depressed populations from subclinical to those in remission concluded that it was most effective when subjects had cognitive impairments before training but were also attentive to the task [56]. Emotional stimuli are infrequently incorporated into CCT training, but this was shown to improve clinical results in one study [11].

**Emotion regulation**
Depressed individuals show increased internal attention to negative representations of the past, which leads to extended rumination, ineffective cognitive reappraisal of negative situations [1], and negative interpretations of affectively ambiguous scenarios [1,2,12]. This latter negative interpretation bias tends to persist despite the revelation of new, positive information [13] and is more pronounced when self-referential prompts are used [14]. Although depression is associated with attention deficits, it can also lead to better concentration on negative self-thought and improved memory for negative information, in contrast to healthy individuals who are more likely to attend to and remember positive information [1]. A possible explanation for this finding is that depressed individuals are more likely to already be in a negative mood, and this congruent context makes negative information easier to remember [1].

One approach for modifying the negative interpretation bias is to train patients to interpret ambiguous words or phrases in a positive way with cognitive bias modification for interpretation (CBM-I). CBM-I has been found to reduce the negative interpretation bias, though sometimes without transfer to other aspects of depression [15].

**External attention**
Tasks that elicit external attention — attention to one’s immediate environment — also capture the negative bias in depression. Although these tasks invoke perceptual processes by presenting negative words and/or images, we do not assume that external attention is the only mechanism at play (e.g., these stimuli could trigger negative memories or rumination). Still, we distinguish these tasks from those discussed above for their separate designs, results, and training attempts. Specifically, we limit our focus to tasks that measure attention towards negative stimuli as well as inhibition of attention away from negative stimuli.

**Spatial attention**
Much of the research in this domain involves variants of the dot-probe task, in which a probe target immediately replaces one of two stimuli that differ in emotion [1,2]. Detection of this probe is faster when it replaces the stimulus being attended than when it replaces the unattended stimulus. In this way, negative bias can be calculated by computing the difference in reaction time to detect a probe that replaces a negative stimulus versus a
positive or neutral stimulus. Such biases are found more reliably when the images are presented for longer durations (e.g. 500–1000 ms [1,2]).

The dot-probe task has been adapted for a form of attention training known as attention bias modification (ABM). In ABM, the probe is placed behind the less negative stimulus more often than the negative stimulus. Such training has been associated with decreases in depression severity [16–18] and improvements in mood [19]. However, other studies found null results and identified potential methodological confounds, including using the same task for training and testing [20]. Indeed, meta-analyses including studies using either ABM or CBM-I show that such training modifies the negative bias in depression inconsistently; CBM-I tends to yield larger effect sizes than those derived from ABM training, both for anxious and depressed samples analyzed together [12] and analyzed separately [21], whatever effect that exists is smaller for depressed than anxious individuals [3,12] and not significant for either clinical population with outliers removed [21], and the number of training sessions is inversely related to study efficacy [21]. Of note, the meta-analyses [3,12,21] included additional ABM paradigms apart from the dot-probe task, but the dot-probe task was the most prevalent [3]. One possible explanation for these mixed results could be how researchers measure depressive symptoms, as the effects of training can be more robust if a stressor precedes the test [12]. Moreover, tasks that rely on spatial probes to measure reaction times at the end of a trial cannot distinguish the roles of specific attention mechanisms [20]; this is better accomplished by studies designed to capture inhibitory processes, discussed next.

**Disengagement**

Where and when people look during natural viewing of emotional images provides a more continuous readout of behavior that can help quantify and distinguish different components of external attention, such as initial selection, sustained attention to relevant stimuli, and disengagement from irrelevant stimuli [20,22–24]. Indeed, eye-tracking studies have shown that depression is associated with slower disengagement from negative stimuli [25] (Figure 2). That is, depressed and non-depressed individuals are equally likely to initially orient towards negative stimuli, but once these stimuli are in the focus of attention, depressed patients struggle to reorient attention elsewhere [24]. A meta-analysis of eye-tracking studies confirmed this finding and further showed that...
the bias for the initial selection component of external attention reflects reduced selection of positive stimuli in depression rather than increased selection of negative stimuli [26].

Other behavioral data point to a problem in inhibiting negative information [27], which may prevent depressed individuals from being able to disengage. One way to measure inhibition is in terms of negative priming — the tendency to respond more slowly to an item after having ignored it on a previous trial when it served as a distractor. (Note: ‘negative’ in negative priming does not have any relation to valence, but is rather used to indicate that the effect is opposite to the typical facilitation by priming.) Larger negative priming indicates successful inhibition and smaller (or nonexistent) negative priming indicates unsuccessful inhibition. Compared to controls, depressed individuals show less negative priming when attending to negatively valenced words and images after having just ignored them as distractors. This suggests a failure to disengage from these stimuli, a tendency that was related to rumination [1]. This disengagement deficit in depression can be found for neutral stimuli as well [1], consistent with reduced inhibition in depression more generally [28]. The attentional differences are not always detrimental: depressed individuals responded faster than controls on a go/no-go task when responding to negative words and ignoring positive words [29].

Given the successful demonstration of negative biases with eye movements, eye-tracking has also been used for attention training, often by guiding overt external attention toward less negative stimuli [30*,31,32]. These paradigms train subjects to look towards more positive stimuli by ending trials only after the participant has fixated on the desired stimulus for a minimum amount of time. Although eye movement training has succeeded in reducing time spent viewing negative stimuli, training has not yet been able to transfer to depression-related symptoms such as mood after a stress-inducing task [31,32]. Another promising training approach involves using eye-tracking to deliver real-time feedback about where a person is looking. In one such study [33**], participants saw positive and negative words and had to unscramble them to form a positive sentence. Their fixations caused the font color to change to either green or red depending on whether they were looking at positive or negative words, respectively. The purpose of this was to help the subjects stay focused on the positive, with the goal of modifying their negative interpretation bias. Eye-tracking feedback might be more effective than merely guiding the eyes away from negative stimuli because incorporating the feedback might engage internal attention processes. Consistent with this, the benefit of eye-based feedback training for reducing the negative attentional bias was mediated by changes in cognitive control [33**].

**Neurofeedback training**

Another class of cutting-edge training approaches does not fit cleanly under internal or external attention, and so we discuss it here separately. These approaches incorporate neural signals into the training, based on findings that fMRI can reveal signatures of the negative attentional bias in depressed individuals not apparent in their behavior [1,2,34–37]. We are not referring to using neural signals as a dependent measure of successful training or symptom change, as has been done in studies that collected fMRI scans before and after behavioral training [16,18,38,39]. Rather, to the extent that specific brain regions are responsible for the negative attentional bias, directly training these regions with neurofeedback from real-time fMRI may induce behavioral improvement. This implicates both internal and external attention because the neurofeedback is based on internal states of the brain related to tasks, goals, and strategies, but it is delivered as a salient visual stimulus in order to externalize these states and influence subsequent processing.

Real-time fMRI neurofeedback is an emerging field, but there is already substantial research demonstrating the feasibility of decoding and altering different aspects of cognition. Monitoring attentional states based purely on neural data has proven possible in healthy subjects. One study found performance differences after trials were triggered in real-time during ‘good’ or ‘bad’ neural attentional states [40]. Another study aimed to reduce the extent to which depressed individuals attended to negative distractors [41], by decoding and providing feedback on how they oriented attention away from negative faces and towards neutral scenes. Other clinically motivated studies trained depressed subjects to regulate specific brain regions in real-time while performing a task. Examples of training paradigms and their respective clinical success include training depressed subjects to: (1) decrease anterior cingulate cortex activity while performing individualized cognitive behavioral therapy strategies on negative autobiographical memories, which predicted self-reported efficacy (post-training) of the cognitive behavioral strategy practiced in real-time [42*]; (2) increase amygdala activity while recalling positive autobiographical memories, which reduced depressive symptoms [43,44**] (Figure 3); and (3) decrease salience network activity while viewing negative images, which lowered self-reported responses to negative stimuli [45].

Neurofeedback based on functional connectivity (i.e. correlation in activity over time between brain regions) is also a promising avenue for attention training, as changes in functional connectivity are associated with reduction in depressive symptoms. For example, upregulating amygdala activity increases connectivity between the left amygdala and left cuneus in a way that predicts reduction of depressive symptoms [46]. More
generally, attention training may increase connectivity within frontal areas and decrease connectivity between frontal and insular areas [39]. Successful treatment has been linked to activity within the insular cortex itself [38] and to connectivity between precuneus and medial frontal gyrus [16]. Even in healthy subjects, training on a neutral flanker task can increase resting-state connectivity between the amygdala and right inferior frontal gyrus [47]. This change in connectivity after cognitive control training could motivate future research to improve cognitive control through neurofeedback that rewards increased connectivity between these regions.

Although fMRI provides unparalleled access to brain states, other neurofeedback technologies have the advantage of being more portable and less costly. For example, the combination of CCT and transcranial direct current stimulation (tDCS) reduces depression symptoms even lasting through a follow-up visit [48]. This benefit was greater than either CCT or tDCS alone, highlighting potentially unique contributions of behavioral and neural attention training. Another approach to translating fMRI neurofeedback to more portable techniques like EEG is to first perform simultaneous fMRI and EEG to derive an EEG correlate of activity only visible to fMRI, such as from the amygdala.
Although not all of the clinical studies mentioned above sought to train attention directly, they offer insight into neural manipulations that could be used to obtain a desired clinical outcome. In the future, internal or external attention could be selectively trained through deliberate design of the task and by targeting specific brain regions. For example, one could use paradigms that load heavily on internal attention, such as reappraising a negative memory to a positive one as in Ref. [42], while targeting neural measures associated with internal attention, such as frontal activity, fronto-limbic connectivity, or decoded emotional states [51]. To train external attention away from negative stimuli, one could present images that vary in terms of image category and valence (e.g. negative faces and neutral scenes as in Ref. [41]), and use a category-based classifier of neural activity to identify, and provide neurofeedback about, how much the negative stimulus is being attended (e.g., how much face evidence there is in the brain).

Conclusions

Depression and other mental health disorders are associated with biases and deficits in attention. Our focus in this review was on the negative attentional bias. We synthesized diverse behavioral and neural findings using modern theoretical frameworks from the study of attention, and then we considered various behavioral and neural training programs that have potential as innovative treatments. Our hope is that this progress will eventually help break the vicious cycle in depression of the bias resulting in increased processing of negative stimuli and thoughts, these experiences worsen mood, which exacerbates the bias and increases processing of negative information, further worsening mood, and so on. In this way, treating the negative attentional bias may lead to a broader reduction in a range of depression symptoms.

Conflict of interest statement

Nothing declared.

Acknowledgements

This work was supported by funding from the John Templeton Foundation (36751) to KAN and NTB, Inel Corporation to KAN and NTB, Canadian Institute for Advanced Research to NTB, and National Institutes of Health (T32 MH065214) to ACM. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


This review provides a very informative account of CCT training and its effect on depression.


Attention and perception


The experimental paradigm described in this paper offers a creative perspective on eye-tracking training. Although these techniques have not yet found success, they deserve to be explored and optimized.


This study offers a unique way to subtly use eye-tracking feedback to make subjects more aware of their internal biases. Eye-tracking feedback is a low-cost and portable way to capture neural processes with high temporal resolution.


This study powerfully combines individualized CBT with neurofeedback. It is noteworthy for this integrative approach and for its demonstration that subjects were more likely to rate the CBT strategies as effective if they were successful in regulating anterior cingulate cortex activity during neurofeedback.


This review provides an excellent introduction to real-time training of depressed individuals based on amygdala activity. The described experiments include the first randomized, double-blind, placebo-controlled design for real-time neurofeedback applied to depression.


This study was the first to report changes in amygdala activity when viewing negative images following extensive practice (18 at-home sessions) of a neutral flanker task, relative to a control group that received an easier version of the flanker task. Even though the participants in this study were healthy, it provides a strong foundation for future work involving depressed patients.


This study used simultaneous EEG and fMRI to identify a scalp EEG correlate of amygdala activity, and then used this EEG signal for...
neurofeedback. This novel approach combines the benefits of fMRI (ability to track subcortical activity) and EEG (cheap, better temporal resolution than fMRI).
