Trinity University
Digital Commons @ Trinity

Psychology Faculty Research | Psychology Department

11-2011

Cognitive Bias Modification: Past Perspectives, Current Findings, and Future Applications

Paula T. Hertel
Trinity University, phertel@trinity.edu

Andrew Mathews
University of California - Davis, ammathews@ucdavis.edu

Follow this and additional works at: http://digitalcommons.trinity.edu/psych_faculty

Part of the Psychology Commons

Publication Details
Perspectives on Psychological Science

Repository Citation
Hertel, P. T., & Mathews, A. (2011). Cognitive bias modification: Past perspectives, current findings, and future applications. *Perspectives on Psychological Science, 6*(6), 521-536.

This Article is brought to you for free and open access by the Psychology Department at Digital Commons @ Trinity. It has been accepted for inclusion in Psychology Faculty Research by an authorized administrator of Digital Commons @ Trinity. For more information, please contact jcostanz@trinity.edu.

aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Cognitive Bias Modification: Past Perspectives, Current Findings, and Future Applications

Perspectives on Psychological Science
6(6) 521–536
© The Author(s) 2011
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1745691611421205
http://pps.sagepub.com
SAGE

Paula T. Hertel[1] and Andrew Mathews[2]
[1]Trinity University and [2]University of California, Davis

Abstract

Research conducted within the general paradigm of cognitive bias modification (CBM) reveals that emotional biases in attention, interpretation, and memory are not merely associated with emotional disorders but contribute to them. After briefly describing research on both emotional biases and their modification, the authors examine similarities between CBM paradigms and older experimental paradigms used in research on learning and memory. The techniques and goals of CBM research are compared with other approaches to understanding cognition–emotion interactions. From a functional perspective, the CBM tradition reminds us to use experimental tools to evaluate assumptions about clinical phenomena and, more generally, about causal relationships between cognitive processing and emotion.

Keywords

cognition, bias, learning (associative), memory, treatment

Folk wisdom has it that some people see their glass as being half full while for others it is half empty. In fact, it is true that people vary in how they interpret ambiguous events and the extent to which they attend to negative or positive information. We also know that these tendencies are associated with susceptibility to emotional distress, such that those who are prone to anxiety or depression are more likely than others to interpret ambiguity in a negative way, attend more to emotionally negative cues, and selectively recall negative information (Mathews & MacLeod, 2005). Like people, animals also vary in their responses to ambiguous events and, concordantly, in their degree of emotional distress. In one study, dogs assessed as being high or low in separation anxiety were trained to expect to find food in a bowl placed in one location but not when it was in another location. When the bowl was put in a position intermediate between these two locations, dogs higher in anxiety were less likely than others to investigate it, apparently assuming that it would be empty (Mendl et al., 2010; for related findings with rats, see Harding, Paul, & Mendl, 2004). Such results illustrate basic connections between emotion and the way in which affectively relevant events are attended, interpreted, and remembered, but until recently, the causal nature of those connections has remained unclear.

This article briefly reviews experiments in the new tradition of cognitive bias modification (CBM). This term (CBM) refers to procedures designed to change particular styles of cognitive processing that are thought to contribute to undesirable emotional reactions or disorders, using systematic practice in an alternative processing style (Koster, Fox, & MacLeod, 2009). CBM experiments have demonstrated that cognitive biases can indeed be modified and that the induced changes influence subsequent emotional reactivity. The approach is not especially novel; it simply involves applying basic principles of experimentation to discover the nature of the connections between cognitive and affective tendencies: A variable believed to affect the putative causal process is manipulated, and the effect on an outcome measure is observed. There is a long tradition in cognitive research of varying affective states—by way of a musical manipulation, for example—to determine whether mood influences some measure of subsequent cognition. In contrast, CBM reverses the direction of cause under investigation to determine whether changes in cognition affect emotional characteristics. The methods we describe are thus quite general in their derivation and applicability and can be used to investigate a wide range of specific questions. Because they are so general, our main purpose is to discuss similarities between CBM paradigms and older, more traditional paradigms used in

Corresponding Author:
Paula T. Hertel, Department of Psychology, Trinity University, 1 Trinity Place, San Antonio, TX 78212
E-mail: phertel@trinity.edu or ammathews@ucdavis.edu

research on learning and memory, such as counterconditioning and transfer-appropriate processing. We also discuss CBM in relation to research on mood-congruent cognition and appraisal theories of emotion. Through such comparisons we hope to emphasize the fundamental principles at work in CBM and to suggest directions for future work. But first things first: We begin by briefly describing the nature of the cognitive biases that have been linked to anxious and depressed states and then turn to the methods that have been developed to modify such biases as well as to the effects of those modifications on emotional reactivity.

Cognitive Biases in Emotional States

Attentional bias

Attention to affective information has commonly been assessed by using designs in which the presence of task-irrelevant but affectively meaningful words or pictures can influence the latency of responses to an intended target. In the emotional Stroop task, for example, naming the color in which words are written tends to be slower when the content of such words is emotionally negative or matches the current concerns of the participant (reviewed by Williams, Mathews, & MacLeod, 1996). More definitive evidence of attentional biases has emerged for search tasks in which participants respond to simple targets (e.g., arrows or letters) that appear in the prior location of either an emotional or neutral cue (e.g., words such as *disease* versus *cabinet*) by pressing a key as quickly as they can. When both emotionally negative and neutral cues are presented simultaneously and replaced a few hundred milliseconds later by a single target in one of the cued locations, anxious individuals are typically faster to respond to targets appearing in locations just vacated by threatening rather than neutral cues, whereas nonanxious control participants show the reverse pattern (MacLeod, Mathews, & Tata, 1986; meta-analysis by Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007). These results have been interpreted as indicating either greater engagement of attention in the vicinity of threat cues by anxious participants, slower disengagement from threat cues, or both. When a single threat cue is presented in one of two possible locations, anxious individuals are particularly slow to detect targets appearing in the other location, suggesting that they have special difficulty in disengaging their attention from threat (Fox, Russo, Bowles, & Dutton, 2001; Yiend & Mathews, 2001). However, studies using endogenous cueing (e.g., neutral or fearful faces with gaze directed to left or right locations) suggest that anxious individuals are both more likely to direct their attention to locations of potential threat—such as toward the direction of fearful gaze—and slower to disengage from such locations (Fox, Mathews, Calder, & Yiend, 2007; Mathews, Fox, Yiend, & Calder, 2003). This conclusion is consistent with clinical observations that people with specific fears orient their gaze toward the possible locations of the object of their fear, as when a person fearful of spiders constantly looks for any sign of spiders in likely hiding places so as to be able to escape if necessary (Rinck, Kwakkenbos, Dotsch, Wigboldus, & Becker, 2010).

Depressed individuals also show similar attentional biases, although these biases seem to operate over a somewhat slower time scale and (unsurprisingly) are directed at stimuli more related to depression than fear (e.g., sad faces; Hankin, Gibb, Abela, & Flory, 2010; Joormann & Gotlib, 2007; but see Mogg, Millar, & Bradley, 2000). In contrast, nondepressed individuals are more likely to attend to positive stimuli. Overall, emotional stimuli tend to capture attention more than do neutral ones, but whereas this tendency in healthy individuals applies as much (and sometimes more) to positive as to negative cues, in those prone to anxiety or depression, attention is more likely to be captured by negative cues, particularly when such cues are related to individual emotional concerns.

Interpretation bias

Relative to nonanxious control groups, anxious individuals are more likely to remember and report ambiguous events in terms of their more threatening meanings. In one early study (Eysenck, Mogg, May, Richards, & Mathews, 1991), participants listened to recorded sentences, some of which were ambiguous and could be interpreted in either a threatening or benign manner (e.g., *the doctor measured little Emma's growth*). Subsequently, they were asked to decide the extent to which related sentences had the same meaning as the original sentences (e.g., *the doctor examined little Emma's tumor* versus *the doctor charted little Emma's height*). The key finding was that nonanxious controls were more likely to endorse nonthreatening interpretations, whereas currently anxious individuals were equally likely to endorse threatening and nonthreatening interpretations. Similar results have emerged from subsequent research. For example, socially anxious and nonanxious participants read descriptions of job interviews and at ambiguous points in the text (when the outcome still remained unclear) were required to make speeded lexical decisions about words related to possible inferences concerning negative or positive outcomes. Nonanxious individuals were consistently faster to endorse words related to positive inferences, whereas socially anxious individuals were equally fast to endorse words matching positive and negative inferences (Hirsch & Mathews, 2000; see also Calvo, Eysenck, & Castillo, 1997; MacLeod & Cohen, 1993; Richards & French, 1992). The critical finding in each of these experiments was an interaction revealing interpretation *differences* between anxious and nonanxious groups, not an outcome indicating that one or the other group was biased in the sense of departing from objective outcome probabilities (that remain unknown). It is clear, however, that higher anxiety levels are associated with a *relatively* greater tendency to perceive the more threatening meaning of emotional ambiguity.

There is also evidence (albeit somewhat less consistent) that depression is associated with negative interpretations of

emotional ambiguity. Depression-related differences were not obtained in an experiment that measured the time to name target words matching negative or benign inferences (Lawson & MacLeod, 1999; also see Bisson & Sears, 2007). However, clear differences did emerge in an experiment using startle modulation to measure the resolution of ambiguity (Lawson, MacLeod, & Hammond, 2002). Another effort (Mogg, Bradbury, & Bradley, 2006) failed to find depression-related differences in reading latencies but found such differences in a spelling task involving the interpretation of spoken homophones (e.g., *wore/war*). Taken together, these results support the contention of Lawson and MacLeod that depression-related differences in interpretation biases may be obscured in reaction time tasks (due to general slowing in depression) but emerge when interpretation is assessed in other ways.

Memory bias

Evidence of negative memory bias in depression has been well established for many years (for reviews, see Mathews & MacLeod, 2005; Matt, Vazquez, & Campbell, 1992). Whereas nondepressed individuals are more likely to remember positive self-descriptive words and happy faces, those prone to depression are relatively more likely to remember negative self-descriptive words and sad faces. Moreover, recent experiments provide evidence that attentional and interpretative biases in depression influence the nature of what is remembered (Hertel & El-Messidi, 2006; Koster, de Raedt, Leyman, & de Lissnyder, 2010; Wells, Beevers, Robison, & Ellis, 2010).

Evidence for anxiety-related biases on memory tests is more mixed, due to some failures to find anxiety-related differences in the recall of threatening versus neutral words (see reviews by MacLeod & Mathews, 2004; Mitte, 2008). The failures have been attributed to ceiling effects (e.g., when words are encoded in relation to the self; Russo et al., 2006) and to the possibility that anxious individuals avoid threatening ideas as a means of emotion regulation (MacLeod & Mathews, 2004). However, as is the case with depression, evidence for memory biases in anxiety is found when other cognitive characteristics of anxiety are taken into account (such as imagery-based distortions in the recall of social scenarios; Hertel, Brozovich, Joormann, & Gotlib, 2008; see Hirsch, Clark, & Mathews, 2006). Anxiety-related memory biases also seem to depend on whether procedures allow for idiosyncratic elaborations. For example, selective recall tends to emerge in experiments in which superficial encoding tasks permit but do not require personally relevant elaboration (MacLeod & Mathews, 2004).

In summary

Compared with the benign or positive processing style that characterizes healthy people, those prone to anxiety and depression are more likely to attend to negative or threatening cues, resolve ambiguity in negative ways, and selectively remember negative events. The critical issues addressed in the work discussed next are whether these cognitive biases can be established experimentally and thereby incur emotional consequences. A fundamental assumption underlying cognitive therapy for emotional disorders and cognitive appraisal theories of emotion is that cognitive characteristics play a part in creating emotional resilience or emotional vulnerability. CBM procedures test this assumption.

Modifying Cognitive Biases

Good experimental control of a phenomenon is achieved if the phenomenon can be simulated and then eliminated. This old saw from the tradition of experimental psychology is made more meaningful by the clinical goal of eliminating biases that characterize emotional distress and disorders. Experiments in the tradition of CBM started by simulating cognitive biases and, in some domains of bias research, continued to the goal of elimination. Unless otherwise specified, the experiments described in the following sections were conducted with nonclinical volunteers; only when participants were selected for special characteristics such as emotional vulnerability or clinical condition do we specify the basis for inclusion.

Attention training

In training experiments, tasks originally developed to measure attention to emotional information are modified in a way designed to promote either attention toward threatening cues (such as words or pictures) or disengagement of attention from them. In the method first described by MacLeod, Rutherford, Campbell, Ebsworth, and Holker (2002), participants are instructed to respond as fast as possible to a neutral target (such as a dot), which is preceded by the display of two words, one threatening and one benign, for 500 ms on either side of fixation. In one group (attend threat), targets nearly always follow in the location of the threat words, while in the other (avoid threat), targets follow in the alternative location. Thus, in the former group, performance is facilitated by attention to the location of threat cues, whereas performance in the latter group is aided by not attending threat cues. Noncontingent probe trials show that attention is indeed deployed as intended, with faster latencies when the target follows in the expected location. However, the critical finding concerns differential emotional responses to a posttraining task involving insoluble problems. The attend-threat group reported more distress than did those trained to avoid threat. In later studies, several sessions of similar training to attend to benign cues were found to reduce reported emotional reactions to a real-life event (moving to another country; See, MacLeod & Bridle, 2009).

In another type of attention task, participants are required to find a target picture in an array of distractors (e.g., a happy face among neutral faces); results have indicated that angry faces can be detected more rapidly than other expressions, especially by socially anxious individuals (Gilboa-Schechtman, Foa, & Amir,

1999; cf. Fox et al., 2000). Dandeneau, Baldwin, Baccus, Sakellaropoulo, and Pruessner (2007) developed this method into a training task by having participants search for one smiling face in a matrix of 16 faces, with the other 15 having rejecting (negative) expressions. In contrast to a control condition in which participants searched for flowers, repeated sessions of training not only reduced rated work stress in telemarketers but also led to lower cortisol levels, an index of emotional distress.

Attention training has also been shown to have beneficial effects for participants suffering from clinical anxiety states, such as Generalized Anxiety Disorder and social phobia (for a meta-analysis, see Hakamata et al., 2010). Practice in responding to targets in the location of words with benign meanings or smiling faces, rather than threatening words or faces, resulted in greater reductions of clinical anxiety than did a control procedure in which targets were presented equally often in either location (Amir, Beard, Burns, & Bomyea, 2009; Amir, Beard, Taylor, et al., 2009; Schmidt, Richey, Buckner, & Timpano, 2009). The finding that most participants in both conditions believed that they had been allocated to a control group, despite substantial improvements in the active condition, potentially argues against accounts based on experimental demand (Amir, Beard, Taylor, et al., 2009), although it is possible that participants were aware of the training contingencies while nevertheless believing that benign training serves as a control for something more emotional. More impressive is the evidence from formal mediation analysis that changes in attention to threat following training mediate the reduction of anxiety in response to stress (e.g., Amir, Weber, Beard, Bomyea, & Taylor, 2008; See et al., 2009).

In summary, repeated practice in avoiding attention to negative cues—such as threatening words or disapproving faces—can reduce subsequent emotional reactions to potentially stressful events, even in highly anxious individuals. The same is true of practice in interpreting ambiguous events in a relatively benign manner.

Interpretation training

As with attention, interpretational style can be modified using variations of tasks originally developed to assess biases in the interpretation of emotionally ambiguous events. In one such modification (of the method used by Eysenck et al., 1991), participants are required to read and imagine ambiguous event descriptions resolved only by the final word, presented as a to-be-completed fragment (i.e., a word with letters missing). For example:

> You have decided to go caving even though you feel nervous about being in an enclosed space. You get to the caves before anyone else arrives. Going deep inside the cave you realize you have completely lost your . . .

This description is followed by the to-be-completed fragment *w_y* (*way*, a negative outcome) or *f_ar* (*fear*, a positive outcome). Participants allocated to practice with either consistently negative or consistently positive resolutions have been subsequently tested with new event descriptions that remain ambiguous, and their interpretations assessed by the extent to which the participants later endorsed either negative or positive statements as having the same meaning as the original descriptions. Compared with the positive group, participants allocated to practicing negative resolutions reported feeling more anxious and made more negative interpretations of the novel test descriptions (Mathews & Mackintosh, 2000). Subsequent experiments showed that effects of training on interpretations persisted until the next day and influenced emotional reactions to viewing videos of real-life accidents. Negatively trained groups reported greater increases in distress than did those who had practiced positive resolutions (Mackintosh, Mathews, Yiend, Ridgeway, & Cook, 2006). More recent experiments using similar methods have produced similar outcomes (e.g., Salemink, van den Hout, & Kindt, 2010).

In an alternative method (used originally by Grey & Mathews, 2000), single ambiguous words (homographs) are followed by to-be-completed word fragments corresponding to associates of the homographs' threatening or benign meanings (e.g., *sink* followed by *dro_n* or *wa_h*). Evidence for the modification of interpretation bias comes from two additional studies using this method (Hoppitt, Mathews, Yiend, & Mackintosh, 2010a; Wilson, MacLeod, Mathews, & Rutherford, 2006). Responses to new test items matching the practiced affective valence were faster than were responses to test items matching the nontrained valence. Furthermore, viewing a subsequent video of accidents elicited more negative emotion in those who had previously practiced accessing negative rather than benign meanings.

As with attention, modifying interpretative style has been shown to produce beneficial changes even in emotionally vulnerable groups. In participants complaining of excessive worry, a single session of practice in selecting the benign meaning of emotionally ambiguous words and descriptions reduced later negative thought intrusions, in contrast to control participants who accessed benign and threatening meanings equally often (Hirsch, Hayes, & Mathews, 2009). This reduction in worry was accompanied by simultaneous improvements on a working memory task, suggesting that changes depended not on increased effort but on increased cognitive control achieved by the reduction of worrying. In a related clinical study, similar training reduced worry in patients suffering from Generalized Anxiety Disorder (Hayes, Hirsch, Krebs, & Mathews, 2010).

Similar beneficial effects of interpretation training have reduced elevated social anxiety (Beard & Amir, 2008). Participants saw either positive or negative words (such as *funny* or *embarrassing*) followed by an ambiguous sentence (*people laugh at something you said*) and were required to decide whether the word and sentence were related. A positive trained group was always given "correct" feedback when they endorsed a positive word and "incorrect" feedback when they

endorsed a negative word, whereas a control group was given noncontingent feedback. Eight sessions of positive training led to greater decreases on a questionnaire assessing social anxiety than in the control group. Moreover, reductions in social anxiety were partially mediated by the extent to which endorsement of positive word–sentence pairs had increased on a separate, noncontingent test of interpretation bias.

Adapted versions of interpretation training have also been found to ameliorate or prevent depressed mood. In one version, participants were trained to make positive interpretations with or without self-imagery, and only the imagery-based training was effective in protecting against negative mood (Holmes, Lang, & Shah, 2009; see also Blackwell & Holmes, 2010). Another adaptation of CBM has been based on previous research showing that rumination (prolonged abstract thinking about the reasons for one's problems) serves to prolong depression; compared with a more abstract, ruminative condition, a new concrete and experiential training condition effectively reduced depressed mood (Watkins, Baeyens, & Read, 2009). In summary, as with attention, practice in interpreting emotional ambiguity in a more positive (or negative) manner influences how novel ambiguous exemplars are interpreted and can also reduce (or increase) negative emotional states or the distress caused by stressful events.

Procedures for changing memory

Cognitive biases revealed on memory tests logically can be produced by biases operating while initially encountering or interpreting the event to be remembered later or by biases operating at the time that memory is tested. Experiments performed to simulate memory biases therefore might involve practice in performing procedures in which the bias is manipulated in any of those ways. Compared with other types of CBM, however, only a few experiments have been performed to establish biased memories, and these experiments mainly have modified interpretation biases in ways that produce distorted memories.

One experiment trained interpretation biases with the procedures established by Mathews and Mackintosh (2000) and found that recall of new ambiguous scenarios presented after training contained distortions that reflected the type of interpretations practiced during training (Tran, Hertel, & Joormann, 2011). In other words, ambiguous statements were remembered as having been nonambiguous in ways congruent with training. A similar experiment simulated situations in which interpretation biases are acquired following the events to be recalled; recall of initial interpretations of the ambiguous scenarios was distorted in line with the training condition experienced subsequently (Salemink, Hertel, & Mackintosh, 2010). This outcome suggests that memory for one's own thoughts can be affected by a bias acquired at a later point.

Two additional experiments used the interpretation-training paradigm combined with a test procedure designed to dissociate controlled and automatic influences on the memory test (Hertel, Vasquez, Benbow, & Hughes, in press; see Jacoby, 1991). These experiments presented a few "critical" scenarios—some resolved negatively and some positively—at the end of a traditional training phase and followed them with a test that produced estimates of controlled recollection of those critical scenarios. The results revealed that experience with benign training scenarios proactively interfered with the recollection of the negative resolutions. This outcome, found with both low and high trait-anxious students, implies that benign interpretation training might be used as a cognitive vaccine against later negative memory biases (see Holmes et al., 2009).

Another means of modifying memory is to reduce the probability of recalling certain kinds of experiences. The tendency to remember and ruminate about negative experiences is worth eliminating, because rumination exacerbates depressive states (see Nolen-Hoeksema, Wisco, & Lyubomirsky, 2008). In a recent experiment designed to promote forgetting on a later test, individuals with major depressive disorder and controls were trained to suppress negative associates to repeatedly encountered benign cues (Joormann, Hertel, LeMoult, & Gotlib, 2009). Depressed participants achieved such forgetting but only when thought substitutes had been used to aid suppression practice (unlike controls who could forget without these aids). This outcome and the previously cited evidence of distorted memory for initial interpretations (Salemink, Hertel, & Mackintosh, 2010) exemplify retroactive effects of practice that modify memory for earlier events. Modifications can therefore influence measures of memory by affecting retrieval processes, not merely initial interpretations.

These five experiments on modifying memory failed to find (or did not assess) evidence of any emotional consequences. However, another study did find that instructions for intentional suppression practice protected against adverse emotional reactions to a subsequent stressful task, compared with a condition in which forgetting was achieved more passively (LeMoult, Hertel, & Joormann, 2010). More generally, the modification of memory bias might be expected to have emotional consequences, because previous experiments have shown that instructions to recall positive autobiographical memories can produce changes in mood states (e.g., Joormann & Siemer, 2004). On the other hand, deliberate reflection on the past is a state of mind that differs from the well-practiced, fluent interpretation of ongoing events that characterizes CBM and is potentially important in producing transfer to emotional reactions during ensuing stress tasks. Perhaps other methods of assessment will better reflect changes in emotion as a consequence of changes in memory.

Whether they are fueled by the pursuit of emotional consequences of memory modification or by other issues, experiments on the modification of memory are ripe for the making. Such experiments can be designed either to alter the contents of memory, to impair memory for the undesirable thought or event, or to improve memory for the desirable thought or event. In the latter regard, efforts to train attention and concentration (e.g.,

MacLean et al., 2010) should also benefit subsequent remembering of emotionally benign events.

In summary

As with modification of attention to threat, practice in the positive resolution of ambiguity can transfer to new ambiguous situations and lead to subsequent changes in affective reactivity to potentially emotional novel events. Similarly, biases induced in interpretation can transfer to tests of memory for the initial ambiguous events and their corresponding interpretations. The evidence discussed so far strongly suggests that cognitive biases are malleable and that modifying them can have consequences for memory and emotional vulnerability.

These conclusions appear to apply across a wide range of conditions and age groups, from the use of interpretation training to reduce social anxiety in children as young as 10 to 11 (Vassilopoulos, Banerjee, & Prantzalou, 2009) to the induction of decreased attention to negative pictures in 60- to 90-year-old adults (Isaacowitz & Choi, 2011). Similarly, the methods have become increasingly varied as they have been applied to different disorders. In one recent application, alcoholic patients were retrained to make avoidant movements to alcohol-related pictures (Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011). Until we learn otherwise, we assume that the same fundamental principles are likely to apply across a wide range of specific applications.

Basic Mechanisms Underlying CBM

Unsurprisingly, the modification of cognitive biases relies on tried and true principles of experimental psychology. Always wary of reinventions, we now call attention to major correspondences between CBM paradigms and older functional paradigms used in research on learning and memory. An advantage in noticing these similarities, beyond parsimony, is that connections to prior research encourage us to reconsider findings from those domains as grist for investigations of biases and their elimination.

Transfer-appropriate processing

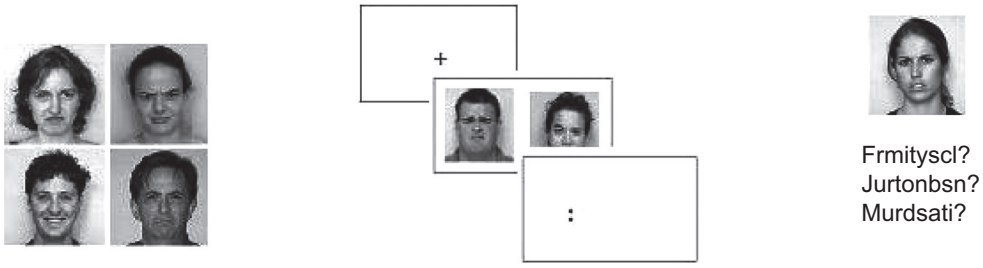
CBM procedures are most certainly related to older experiments on transfer of learning, popular in the mid-twentieth century (see Ellis, 1965), as well as more current approaches to transfer in concept learning and problem solving (e.g., Nokes, 2009). Researchers such as Ellis wrote about near and far transfer to represent the degree of overlap between training and transfer tasks. Although the dimension of near to far transfer should be viewed as a continuum, most CBM experiments are near-transfer experiments because, as illustrated in Figure 1, the situations presented and procedures invoked during training are similar to those in the transfer phase. Some stressful transfer tasks (e.g., emotional response to viewing a video of dangerous accidents; Wilson et al., 2006) can be thought of as far-transfer tasks; the same cognitive procedures are assumed to be involved but the contexts differ substantially from those presented during training.

Near-transfer tasks in CBM have been used to explore the extent to which training in one type of attention or interpretation task generalizes to other tasks with similar processing requirements. Practice in the interpretation of homographs by completing word fragments thus generalizes to lexical decision speed in judging associates of ambiguous primes (Grey & Mathews, 2000; Hoppitt et al., 2010a; Wilson et al., 2006); practice in judging semantic similarity between such primes and related (benign or negative) words generalizes somewhat farther to the content of mental images cued by novel homographs (Hertel, Mathews, Peterson, & Kintner, 2003). Similarly, training via searching for a smiling face in a matrix of negative expressions reduces interference from rejection-related words in a modified Stroop task (Dandeneau & Baldwin, 2004). Near-transfer effects can sometimes fail, however, when the type of response required or the content domain differs across training and test (Salemink, van den Hout, et al., 2010; Salemink, van den Hout, & Kindt, 2007). Far-transfer effects occur in the critical experiments used to establish causal links between cognitive processing bias and emotional reactivity. In these experiments, training selective attention or interpretation using simple word stimuli can influence later emotional reaction to apparently very different experiences, such as failing to solve difficult anagrams or viewing accident videos (Mackintosh et al., 2006; MacLeod et al., 2002; Wilson et al., 2006).

In the field of memory research, the concept of transfer has proved to be useful across the decades and into this century, mainly as a theoretical framework called transfer-appropriate processing (TAP; Morris, Bransford, & Franks, 1977). The initial experiments in the TAP framework revealed that performance on memory tests was facilitated when the cognitive procedures invoked on the test replicated those at work during initial encounters—for example, when conceptual cues for recall, not perceptual cues, followed an initial semantic elaboration task.

TAP experiments have been performed in a variety of contexts for learning and testing. One of the more applicable examples is found in a series of experiments on spontaneous transfer in problem solving (Needham & Begg, 1991). At the time of this work, few examples could be found of uninstructed transfer of solutions to logic-based word problems. The training problems were typically taught as examples to be studied and learned, whereas the transfer phase consisted of new problems to be solved. Good levels of transfer of solution methods to new problems was obtained when participants were told before the presentation of target problems to think back to analogous prior problems but not when the analogies were left implicit. This was a distressing state of affairs for applications to education, and we note that a parallel exists in therapeutic contexts. Just as the teacher is not present in later classes to remind the student of analogies learned in her class, the

Attention Training	Near Transfer	Far Transfer (Challenge)
Latency to Find A Smiling Face Among Rejecting Faces.	Latency to Identify Target Replacing Either a Negative or Positive Face.	Rated Emotion Follows Attempt to Solve Anagrams While Viewing Disapproving Face.



Interpretation Training	Near Transfer	Far Transfer (Stress)
Latency to Complete Fragment Related to the Threatening or Benign Meaning of the Just-Presented Homograph.	Latency to Choose the More Related of two Words Following a Homograph on the Previous Screen.	Rated Mood Before and After a Video of a Series of Actual Accidents.

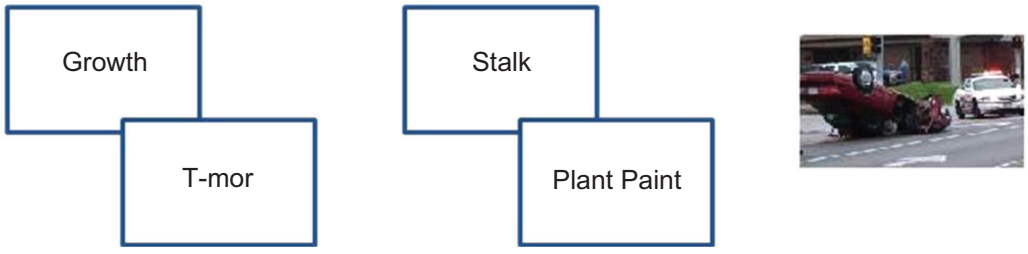


Fig. 1. Examples of phases from an attention-training paradigm (Dandeneau & Baldwin, 2009) and an interpretation-training paradigm (Wilson et al., 2006). Training phases typically consist of 80–100 trials. The dependent measures are listed in the table, and examples of computer displays are shown below each.

clinician is not present during real-world "tests" to remind the client to think back to their sessions. By applying the TAP framework, however, Needham and Begg invented the proper training analog whereby participants were instructed not to learn the training problems but instead to solve them; under those conditions, spontaneous transfer to solving new problems was obtained. Conversely, studying the training problems with the goal of remembering them transferred well to target solutions when participants were actually instructed to remember specific training problems at the time of the test.

Interpretation training sessions in CBM experiments also establish the conditions for spontaneous transfer. Far-transfer effects following interpretation training depend on the active selection of emotional meanings during training rather than simple exposure to them. For example, emotional reactivity differences have been observed after training in which participants resolved emotional meanings of ambiguous words or texts for themselves but not when they read already disambiguated versions (Hoppitt et al. 2010a; Hoppitt, Mathews, Yiend, & Mackintosh, 2010b). Even larger effects would be expected if CBM participants were gradually encouraged to resolve training scenarios in the absence of fragments. More generally, transfer in CBM experiments seems to be successful to the extent that the training and transfer "problems" invoke the same cognitive procedures. Successful transfer to later emotional tasks thus depends on practice in actively selecting emotional meanings during training because elicitation of affect at test depends on similar active involvement. Nevertheless, even when this condition is met, far transfer to emotional challenges is not always obtained (e.g., Salemink et al., 2007; Teachman & Addison, 2008). A major goal for CBM researchers is to find ways to establish robust far-transfer effects in order to sidestep the need for individualized training in stressful situations.

Desirable difficulties

Another goal is to use modification procedures that last. One sure way to promote better memory on delayed tests is to distribute practice across sessions instead of massing the same amount of practice into one session (see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). In a transfer-appropriate fashion, having something come to mind after a period of not thinking about it facilitates the same retrieval process on subsequent attempts. In this regard, a few CBM experiments have taken advantage of the benefits of distributed practice when using multiple training sessions with clinical groups (e.g., Amir et al., 2009).

Distributed practice is one of several operations referred to as a desirable difficulty by researchers who attempt to apply memory principles in educational settings (Bjork, 1994). Students and instructors alike tend to prefer "easy" instructional techniques that lead to rapid changes in performance (e.g., Kornell & Bjork, 2008); the benefits of practice are quickly observed, and students believe they learn better. Such preferences are probably also held by some therapists and their clients, indeed by the culture at large. Media emphasize ways of making learning easy and fun. Unfortunately, students often forget what they have learned in college classes once the test has been taken, and anxious or depressed individuals easily revert to well-practiced biases outside the therapeutic context. Therefore, the search for desirable difficulties in CBM research is worthwhile.

In addition to distributed practice, other desirable difficulties identified by applied memory research include the initial generation of the information to be remembered, in comparison to merely reading it (Bjork, 1994). As we have discussed previously, this procedure likely works because it is transfer appropriate to the demands of the "test." (Transfer-inappropriate difficulty is undesirable.) A very similar prescription is to take advantage of the testing effect (Roediger & Karpicke, 2006); a series of tests promotes better performance on a final test, compared with a series of reexposures to the to-be-remembered materials. Similarly, in clinical situations, therapists sometimes try to simulate stressful distractions of real life so that clients can practice the prescribed coping under difficult conditions. Establishing variable conditions during practice is also a desirable difficulty, because variable conditions promote good performance in new situations (see Bjork, 1994; McDaniel & Callender, 2010). Anxious and depressed individuals are faced with a variety of ambiguous situations in real life. When one is not sure what conditions will hold during real-life tests, the wise course is to vary conditions during training to optimize transfer (e.g., Mackintosh et al., 2006).

In summary, the main criteria regarding cognitive procedures to be used in modifying cognitive biases are their compatibility with those occurring during the challenges that life presents and their tendency to promote changes in the long term. These are also the lessons learned from much older traditions of associative learning.

Associative learning

With roots in the infancy of experimental psychology, contemporary learning theory provides useful perspectives on the etiology and maintenance of emotional disorders, anxiety disorders in particular (see Mineka & Zinbarg, 2006). Fear-related Pavlovian or operant contingencies are established through direct or vicarious conditioning and aided by human tendencies to be evolutionarily prepared for them. Any attempt to change emotional reactions therefore should be understandable in terms of basic principles of associative learning, such as extinction and counterconditioning. It is important to note that these methods of behavior change do not cause unlearning of the fear contingencies but instead establish new contingencies that strongly rely on their context (Bouton, 2000). Lapse and relapse, in which the original contingencies dominate to guide behavior, tend to occur if features of the context revert to earlier conditions (the renewal effect) or if the original significant outcomes occur noncontingently (reinstatement). Spontaneous recovery of the original response is also likely to occur when fear-evoking cues are presented following a period of time subsequent to extinction or counterconditioning trials (i.e., following time after therapy sessions have ended).

Do similar dangers of lapse and relapse of old habits characterize CBM procedures? In interpretation retraining, contingencies between ambiguous situations and their preexperimental resolutions are, in effect, counterconditioned. Consider anxious participants, for whom the signal is an ambiguous situation harboring the possibility of threat; the original significant outcome (the unconditioned stimulus) is the threatening resolution, and the new outcome is a benign resolution. Bouton (2000) actually described the signal in typical extinction or counterconditioning paradigms as having two possible meanings, one old and one new; the current meaning of the ambiguous signal is determined by the exteroceptive, interoceptive, or temporal context. At first, when the context is changed in the animal laboratory (e.g., from one apparatus to the next), behavior generalizes (or in theoretical terms, the expectancy of the original outcome holds), much like when human participants begin a series of training trials and reveal their typical resolutions of ambiguity, trained by real life. Then counterconditioning trials ensue, and we can understand them from a Pavlovian perspective of an ambiguous (negative or positive) cue predicting the targeted positive resolution. (See Clerkin & Teachman, 2010, for a similar application of conditioning principles to CBM.) Alternatively, from an operant perspective, an ambiguous cue can set the occasion for the reinforcement of target resolutions. Most forms of CBM include reinforcement of some kind, in that tasks are designed to reward the desired response (e.g., attending to a benign stimulus or meaning), either because doing so makes performance easier and more fluent or because overt feedback is provided (e.g., a "correct" signal). As we also learn from research on desirable difficulties, variable conditions of

training promote transfer beyond the conditioning episodes, as do temporally spaced training sessions (Bjork, 1994; Bouton, 2000).

If the results from animal conditioning experiments apply to CBM paradigms, however, signals presented in contexts that differ from the CBM context should lead to robust renewal of the initial response, and occasional experiences with negative events should reinstate it. In part, renewal, reinstatement, and spontaneous recovery might be responsible for occasional difficulties in finding evidence of far transfer after CBM. Researchers with interests in therapeutic applications have tried to overcome that difficulty in ways that are consistent with Bouton's (2000) analysis. One method is to promote retrieval of the context for new learning—the training phase. Methods for reminding participants about training experiences as they encounter adverse events should enhance the chance of generalization or transfer, as should the use of ecologically relevant episodes during training. In related investigations of exposure therapy with phobic participants, instructions to think back to the therapy session during behavioral testing were found to improve transfer to a new context (Mystkowski, Craske, Echiverri, & Labus, 2006). Reminders can also be offered by others; in one anecdote reported by Brosan, Hoppitt, Shelfer, Sillence, and Mackintosh (2011) after an anxious patient described his CBM session to his wife, she began saying "beep" (the error signal) to him whenever she saw him begin worrying.

In short, the literatures concerning conditioning and memory principles converge in ways that are relevant to treatment with CBM. Applications of CBM should include methods for overcoming renewal, reinstatement, and spontaneous recovery of the preexperimental bias. For example, the foregoing discussion suggests that training should incorporate situational cues that typically provoke the deployment of negative biases in real life. Subsequent encounters with such real events should then provide automatic reminders of the newly acquired adaptive response. More generally, awareness of this analogy between CBM and conditioning paradigms reminds us about the values inherent in functional or behavioral approaches to understanding what we commonly consider to be mental phenomena (see De Houwer, 2011).

Considerations of automaticity versus cognitive control

Although we have recommended the use of research on desirable difficulties as a model for how to achieve transfer in the modification of cognitive biases, we acknowledge that experiments concerning desirable difficulties largely involve situations in which individuals explicitly attempt to remember, and CBM transfer tasks usually do not use explicit instructions to think back to training. Alternatively, the analogy to conditioning paradigms promotes the idea that change is achieved implicitly through the modification of cognitive habits, without conscious reference to the training phase. In attempting to define CBM paradigms, for example, Koster et al. (2009, p. 3) suggested that "[c]ommonly, it has been the case that the targeted bias represents a pattern of processing selectivity that appears to operate automatically, in the sense of proceeding swiftly without intention, and so is not readily amenable to volitional control." However, just as memory tests are not "process pure" (Jacoby, 1991), training sessions likely operate in habitual ways to influence current attention and interpretation but also consciously come to mind in a recollective sense. In fact, participants sometimes report consciously thinking back to specific training scenarios as they perform real-life transfer tasks (Blackwell & Holmes, 2010). Similarly, training can be more effective following explicit instructions about the contingency between cues and targets (Krebs, Hirsch, & Mathews, 2010). On the other hand, postexperimental interviews usually fail to reveal awareness of the training manipulation in the sense that participants later deny seeing any causal connection between the training procedure and their own later reactions in challenging situations (e.g., Amir, Beard, Taylor, et al., 2009; Hirsch et al., 2009). These reports of lack of awareness might be more convincing if we could rely on the accuracy of self-reports about recollection during transfer that are made subsequently, but self-reports about cognitive procedures are usually not reliable. In short, an important direction for bias-modification research is the exploration of the extent to which training experiences establish new habits or provide the basis for controlled recollection, as well as the contexts under which automatic or controlled processes dominate in producing good transfer.

What do we know about the involvement of controlled procedures in training paradigms? First, practice in deliberately suppressing (i.e., controlling "against") the retrieval of learned negative responses clearly impairs later deliberate recall, particularly when thought substitutes are used to help participants not think about the target words to be suppressed (Joormann et al., 2009). This is a finding that challenges popular assumptions about the inevitability of rebound following attempts at thought suppression. The success of suppression-induced forgetting, replicated in many experiments using the think/no-think paradigm (see Anderson & Huddleston, in press), likely results from the many trials of practicing suppression in response to specific cues, whereas other methods offer less well-controlled and repetitive means.

Deliberate recall has also been instructed during the transfer phase of a more typical paradigm for interpretation training (Hertel et al., in press). Participants were instructed to complete some transfer scenarios in the same way as they had completed contextually similar prior scenarios and to complete others differently. For example, if a training scenario described being on a plane to Florida, the analogous transfer scenario might describe being on a plane to Spain. When instructed to respond as before, the correct response to the transfer scenario was to complete it in the same way (i.e., benign or negative) as the related original or in the converse manner if instructed to respond differently. Estimates of

controlled recollection were calculated as the difference between the proportion of correct completions under instructions to respond in the same way and the proportion of incorrect completions under instructions to respond differently (see Jacoby, 1991). The main finding was that benign training proactively interfered with the controlled recollection of negative interpretations. This finding does not demonstrate training-congruent facilitation of control, because it represents deficient recollection of test scenarios previously resolved in a training-incongruent manner, not enhanced recollection for those resolved in the trained direction. The lack of facilitation might reflect what we learned from research on spontaneous transfer in problem solving—that problem-oriented training does not provide the best transfer to a memory-oriented test phase (Needham & Begg, 1991). The process-dissociation experiments (Hertel et al., in press) also failed to find evidence for effects of training on automatic components of transfer, but this outcome should be interpreted in light of the fact that CBM transfer tasks do not usually call attention to the connections between training and transfer, as is required by process-dissociation procedures at test. Other methods of assessing automatic influences might tell a different story about automatic components of transfer.

As yet there has been little neuroimaging research that could cast light on the contribution of cognitive control to bias modification. Evidence for greater activation of the lateral frontal cortex during test trials that required attention to be directed differently from that previously practiced was taken as evidence that such activation mediates training effects (Browning, Holmes, Murphy, Goodwin, & Harmer, 2010). Because the lateral frontal cortex is implicated in intentional control, this finding might argue against exclusively automatic effects. However, the effect was observed during violations of the trained rule during test trials (i.e., when target detection required attention to the alternate location to that induced during training), so it may be that intentional control is implemented only when participants have to voluntarily override the practiced responses. A prediction that remains to be tested is whether a similar pattern of activation during training that counteracts preexperimental bias would decrease toward the end of the training trials and thereby constitutes evidence for a newly trained habit.

In summary, despite the apparent importance of this issue, it is simply too soon to know to what extent and in what ways controlled and automatic processes are involved in CBM. Resolution will require care in specifying the sense in which procedures should be conceived as controlled or automatic (see Moors & De Houwer, 2006). Clearly, experimental control of practice is ultimately responsible for transfer effects in CBM, so the issues concern the degree to which and manner in which individuals' use of self-initiated conflict resolution or deliberate search processes (frontally mediated procedures) characterizes, benefits, or impairs transfer and whether those effects are restricted to initial stages of bias modification. The extreme position that performance on near- or far-transfer tasks is intentionally made to be consistent with training is refuted by participants' denial of connections between training and transfer tasks. Moreover, awareness of connections to the past can be epiphenomenal to essentially automatic responding. As in learning a new motor skill like riding a bicycle or hitting a tennis ball, performance can come under the control of practice without any awareness of the underlying automatic components that must be recruited or inhibited. If CBM is similar to motor learning in this respect, then fundamental processes influencing ease of access to emotional meanings may be modified during practice, yet produce changes that are typically attributed to controlled operations such as the intent to reinterpret. In short, people do not always judge the cause of their actions accurately, even while experiencing intent.

Relation to Other Research on Cognition–Emotion Interactions

In this section, we compare the procedures and outcomes of CBM experiments to those found in other research traditions concerned with cognition and emotion interactions. In some cases, the understanding of CBM findings can be informed by reference to more established research domains, but in other cases we argue that CBM methods offer more nuanced tools that can be used to research questions relevant to these other areas.

Mood and cognition

CBM experiments are not the first experiments on affect and cognition to examine the direction of cause. Various methods associated with other research traditions have sought to establish emotion as both the cause and effect of thoughts and memories (see Parrott & Hertel, 1999). Like CBM research, some of this research has focused on emotional states that result from particular styles of thinking. For example, the manipulation of ruminative versus distracting thoughts has produced changes in depressed mood (e.g., Fennell & Teasdale, 1984; Nolen-Hoeksema & Morrow, 1993). Manipulations of rumination, although not devised for the purpose of establishing or eliminating specific cognitive biases, are quite similar to CBM manipulations in that they also have been shown to produce changes in performance on subsequent cognitive tasks, some of which are good examples of far transfer (e.g., Hertel & El-Messidi, 2006; Joormann & Siemer, 2004; Lyubomirsky, Caldwell, & Nolen-Hoeksema, 1998).

The opposite direction of causality in cognition–emotion interactions has served as the focus of a much larger literature on the effects of mood manipulations. Among other ways, mood ostensibly has been manipulated through hypnosis, self statements, success or failure, films, pictures, and music (see Parrott & Hertel, 1999), and consequences for performance on tests of attention, judgment, and memory have been examined. In our view, many of these experiments are also similar to CBM experiments, because the manipulations are essentially

cognitive. They are viewed as mood manipulations primarily because researchers perform manipulation checks to determine if mood has been changed in the expected directions, but the causal status of mood is typically not established. The classic example is the ubiquitous Velten mood-induction procedure (Velten, 1968), in which participants read and think about a series of statements about oneself or the world that are emotionally positive, neutral, or negative. This manipulation overtly changes the nature of thoughts in ways that have subsequent cognitive and self-reported emotional consequences, much as bias training transfers to other cognitive tasks and affects subsequent emotional reactions.

In all traditions (CBM and earlier cognition–emotion experiments), we cannot always be certain that the direction of cause conforms in simple ways to the focus of our manipulation. Manipulations of cognitive procedures might masquerade as mood manipulations or vice versa. Thoughts can have immediate emotional consequences that in turn direct the focus of other thoughts. Our inferences about cause are only as tight as the experimental design permits. In that regard, however, it is important to note that several CBM experiments have not produced changes in mood states immediately following training. In fact, although the scenario-based method of interpretation training (Mathews & Mackintosh, 2000) can lead to immediate mood changes, most other methods (involving interpretation or attention to single words) do not, as assessed by self-rating or questionnaire measures (e.g., Hoppitt et al., 2010a, 2010b; MacLeod et al., 2002; Wilson et al., 2006). Despite this lack of mood induction by training, differences in emotional vulnerability can still be revealed later, in response to a potentially threatening or distressing event. Furthermore, a direct comparison between a traditional (musical) mood induction and interpretation training showed that the latter resulted in congruent changes on the scrambled sentence test (Rude, Valdez, Odom, & Ebrahimi, 2003), but there were no such changes on the same test following mood manipulation (Standage, Ashwin, & Fox, 2010). Another experiment in which direction of training and mood induction were independently manipulated showed that posttraining mood did not modify responses to a subsequent transfer test to assess interpretation bias (Salemink & van den Hout, 2010). In short, CBM cannot simply be seen as a variant on mood manipulations. Typically, mood manipulations do not systematically guide attention or the interpretation of ambiguity, and they are not designed to promote transfer to similar situations.

In a more general sense, cognitive bias manipulations differ from previous manipulations in the specificity of their methods and goals. The previous methods were designed to induce general cognitive or affective states (such as rumination or sadness) believed to have general consequences (negative feelings or memories) by virtue of some theoretical mechanism such as activation in a semantic network (e.g., Bower, 1981). CBM methods target specific cognitive procedures with the potential to emerge in a variety of circumstances (such as attending to a threatening thought or event or interpreting ambiguity in a negative light) and that are practiced in ways that should transfer to similar new situations. In our view, the compatibility of CBM with functional analyses derived from associative learning and transfer-appropriate-processing frameworks sidesteps the global assumptions that characterized early research on cognition and emotion (also see MacLeod, Koster, & Fox, 2009).

Appraisal theory

Appraisal theorists have long proposed that emotions arise from a cognitive process in which individuals evaluate the personal implications of an event, including their ability to control or cope with it, and that variations in emotion depend primarily on the outcome of this process (e.g., Roseman, 1991; Smith & Lazarus, 1993). The CBM approach clearly shares the common goal of investigating the assumed role played by cognitive processes in emotion, yet the methods of investigation differ. Appraisal theories have generally been tested by showing that the self-reported relevance of an event is a good predictor of emotional response and not by modifying processes assumed to contribute to appraisal but likely inaccessible to conscious report, as in CBM experiments. Application of modification methods might thus prove useful in testing tenets of appraisal theory. Conversely, the extensive evidence of individual differences found in appraisal research should usefully inform future CBM experiments.

An example of the first direction of possible cross-application is provided by Schartau, Dalgleish, and Dunn (2009), who showed that practice in adopting a benign appraisal style during exposure to aversive pictures resulted in a later reduction of negative emotion during the viewing of a video depicting serious accidents. Further experiments would be required to provide more precise information on the specific processes underlying such instructed appraisal. Concerning the reverse direction of application, it might be possible to understand some failures to find far-transfer effects in CBM experiments by considering individual differences in appraisals. For example, the absence of any benefit from training positive social interpretations on later reactions to failing a cognitive test (Salemink et al., 2007) might in retrospect be understood in terms of the different appraisals associated with affiliative and achievement concerns (Smith & Pope, 1992). Design of future far-transfer experiments might thus profit from knowing about the emotional consequences of variations in appraisal style across individuals and events. Conversely, CBM methods should provide a useful tool to test specific causal assumptions of appraisal theory in a more analytic way.

Relation to Cognitive Behavior Therapy

CBM applications to clinical disorders overlap to some extent with cognitive behavior therapy (CBT). Both involve direct attempts to modify cognitive processes believed to underlie emotional distress, and both employ systematic exposure to the events that trigger distress. In CBT, exposure to evocative

events is designed to reduce emotional reactions via extinction (sometimes referred to as habituation) or to collect evidence against false beliefs by showing that anticipated catastrophic consequences do not occur. Indeed, it is often thought that attempts to avoid accessing fear-evoking representations serve to maintain anxiety, because stored emotional memories must be fully activated before they can be changed (Foa & Kozak, 1986). This hypothesis sometimes leads clinicians to predict that avoidance of attention to threats can worsen emotional symptoms. In CBM, by contrast, extinction due to exposure is not thought to be the agent of change; in fact, training and control conditions sometimes involve a similar extent of exposure to potential threat cues. Instead, the critical difference between training and control conditions is how such evocative events are processed (i.e., whether potential threat cues are to be attended or not, or how they are interpreted or recalled). The effectiveness of CBM thus suggests that fear-evoking events do not need to be fully realized for change to occur, as some clinical researchers have assumed, because training to avoid threats—in favor of alternative meanings or events—can be beneficial. This conclusion suggests possible alternative explanations for the effectiveness of CBT, such as alterations in the relative ease of retrieving positive versus negative emotional memories (see Brewin, 2006). Direct empirical comparisons between CBT and CBM perspectives on fear, investigating the extent to which they have overlapping or differential effects, would thus constitute a worthwhile direction for future research.

To be sure, CBT also targets inappropriate cognitive content, but here the usual means of change is to directly challenge the reportable thoughts believed to maintain distress. In contrast, CBM involves repeated trials that target processes typically not available to conscious report, and as we have argued earlier, participants are not necessarily aware of the resultant changes. Nonetheless, as CBM moves into therapy settings, this last distinction might become less pronounced, with therapists giving their clients more explicit guidance about what is required (Krebs et al., 2010). (In this regard, we should consider the possibility that explicit instructions for change might sidestep processes that promote longer lasting effects.) Similarly, as CBT embraces the use of computer-controlled methods (Andrews, Cuijpers, Craske, McEvoy, & Titov, 2010), possibilities for integration are likely to increase.

Conclusion

From the outset, CBM research has been motivated and guided by hypotheses emerging from clinical research and practice, such as hypotheses concerning the causal role for cognitive processes in emotional disorders postulated by pioneers such as Beck (1976). As in other experimental approaches, hypothesis testing has forced CBM researchers to be more precise about putative causes when designing modification methods. Causal hypotheses are then tested in what has become known as experimental-psychopathology research, in which symptoms of disorders may be produced or reversed. One example is provided by the work cited earlier showing that changing the style of thinking associated with rumination reduces depressed mood (Watkins et al., 2009). In another example, the hypothesis that depression is maintained by the tendency to attribute negative outcomes to the self and positive outcomes elsewhere was confirmed by experimentally inducing this attributional style in one group and the opposite style in another (Peters, Constans, & Mathews, in press). CBM research also allows the investigation of competing hypotheses, such as whether pathological worry is exacerbated by a failure to disengage attention or by enhanced engagement with threatening content (Hirsch et al., 2011). Thus, starting from the general idea that the biases characteristic of emotional disorders may contribute to maintaining them, CBM researchers develop methods for testing specific hypotheses about the conditions under which cognitive processes exacerbate or reduce emotional symptoms, and then the results of those tests can be used to guide the development of new treatments.

Beyond strictly clinical concerns, research paradigms gathered together under the rubric of cognitive bias modification are nothing more and nothing less than tools to test assumptions, often initially unspecified, about interactions between cognitive and emotional processes. If we think we understand a phenomenon, we design a procedure to simulate it with naive individuals by bringing the putative processes under experimental control. Transfer-appropriate and contextually sensitive operations constitute basic rules of engagement. If we understand it well enough, the phenomenon in question can be not only simulated but also manipulated to produce changes in mental and emotional well-being.

Acknowledgments

The authors contributed equally to the article.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

References

Amir, N., Beard, C., Burns, M., & Bomyea, J. (2009). Attention modification program in individuals with generalized anxiety disorder. *Journal of Abnormal Psychology, 118*, 28–33.

Amir, N., Beard, C., Taylor, C.T., Klumpp, H., Elias, J., & Chen, X. (2009). Attention training in individuals with generalized social phobia: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 77*, 961–973.

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

Anderson, M.C., & Huddleston, E. (in press). Towards a cognitive and neurobiological model of motivated forgetting. In R.F. Belli (Ed.), *True and false recovered memories: Toward a reconciliation of the debate. Nebraska Symposium on Motivation* (Vol. 58). New York, NY: Springer.

Andrews, G., Cuijpers, P., Craske, M.G., McEvoy, P., & Titov, N. (2010). Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: A meta-analysis. *PLoS ONE, 5*(10), e13196. doi:10.1371/journal.pone.0013196

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M.J., & van IJzendoorn, M.H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin, 133*, 1–24.

Beard, C., & Amir, N. (2008). A multi-session interpretation modification program: Changes in interpretation and social anxiety. *Behaviour Research and Therapy, 46*, 1135–1141.

Beck, A.T. (1976). *Cognitive therapy and the emotional disorder*. New York, NY: International Universities Press.

Bisson, M.A.S., & Sears, C.R. (2007). The effect of depressed mood on the interpretation of ambiguity, with and without negative mood induction. *Cognition & Emotion, 21*, 614–645.

Bjork, R.A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalf & A. Shimamura (Eds.), *Metamemory: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.

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

Bouton, M. (2000). A learning theory perspective on lapse, relapse, and the maintenance of behavior change. *Health Psychology, 19*, 57–63.

Bower, G.H. (1981). Mood and memory. *American Psychologist, 36*, 129–148.

Brewin, C.R. (2006). Understanding cognitive behaviour therapy: A retrieval competition account. *Behaviour Research and Therapy, 44*, 765–784.

Brosan, L., Hoppitt, L., Shelfer, L., Sillence, A., & Mackintosh, B. (2011). Cognitive bias modification for attention and interpretation reduces trait and state anxiety in anxious patients referred to an out-patient service. *Journal of Behavior Therapy and Experimental Psychiatry, 42*, 258–264. doi:10.1016/j.jbtep.2010.12.006

Browning, M., Holmes, E.A., Murphy, S.E., Goodwin, G.M., & Harmer, C.J. (2010). Lateral prefrontal cortex mediates the cognitive modification of attentional bias. *Biological Psychiatry, 67*, 919–925.

Calvo, M.G., Eysenck, M.W., & Castillo, D.M. (1997). Interpretation bias in test anxiety: The time course of predictive inferences. *Cognition & Emotion, 11*, 43–63.

Cepeda, N.J., Pashler, H., Vul, E., Wixted, J.T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*, 354–380.

Clerkin, E.M., & Teachman, B.A. (2010). Training implicit social anxiety associations: An experimental intervention. *Journal of Anxiety Disorders, 24*, 300–308.

Dandeneau, S.D., & Baldwin, M.W. (2004). The inhibition of socially rejecting information among people with high versus low self-esteem: The role of attentional bias and the effects of bias reduction training. *Journal of Social and Clinical Psychology, 23*, 584–602.

Dandeneau, S.D., & Baldwin, M.W. (2009). The buffering effects of rejection-inhibiting attentional training on social and performance threat among adult students. *Contemporary Educational Psychology, 34*, 42–50.

Dandeneau, S.D., Baldwin, M.W., Baccus, J.R., Sakellaropoulo, M., & Pruessner, J.C. (2007). Cutting stress off at the pass: Reducing vigilance and responsiveness to social threat by manipulating attention. *Journal of Personality and Social Psychology, 93*, 651–666.

De Houwer, J. (2011). Why the cognitive approach in psychology would profit from a functional approach and vice versa. *Perspectives on Psychological Science, 6*, 202–209. doi:10.1177/1745691611400238

Ellis, H.C. (1965). *The transfer of learning*. New York, NY: Macmillan.

Eysenck, M., Mogg, K., May, J., Richards, A., & Mathews, A. (1991). Bias in interpretation of ambiguous sentences related to threat in anxiety. *Journal of Abnormal Psychology, 100*, 144–150.

Fennell, M.J.V., & Teasdale, J. (1984). Effects of distraction on thinking and affect in depressed patients. *British Journal of Clinical Psychology, 23*, 65–66.

Foa, E.G., & Kozak, M.J. (1986). Emotional processing and fear: Exposure to corrective information. *Psychological Bulletin, 99*, 20–35.

Fox, E., Lester, V., Russo, R., Bowles, R.J., Pichler, A., & Dutton, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition & Emotion, 14*, 61–69.

Fox, E., Mathews, A., Calder, A., & Yiend, J. (2007). Anxiety and sensitivity to gaze direction in emotionally expressive faces. *Emotion, 7*, 478–486.

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General, 130*, 681–700.

Gilboa-Schechtman, E., Foa, E.B., & Amir, N. (1999). Attentional biases for facial expressions in social phobia: The face-in-the-crowd paradigm. *Cognition & Emotion, 13*, 305–318.

Grey, S., & Mathews, A. (2000). Effects of training on interpretation of emotional ambiguity. *Quarterly Journal of Experimental Psychology, 53*, 1143–1162.

Hakamata, Y., Lissek, S., Bar-Haim, Y., Britton, J.C., Fox, N.A., Leibenluft, E., . . . Pine, D.S. (2010). Attention bias modification treatment: A meta-analysis toward the establishment of novel treatment for anxiety. *Biological Psychiatry, 68*, 982–990.

Hankin, B.L., Gibb, B.E., Abela, J.R.Z., & Flory, K. (2010). Selective attention to affective stimuli and clinical depression among youths: Role of anxiety and specificity of emotion. *Journal of Abnormal Psychology, 119*, 491–501.

Harding, E.J., Paul, E.S., & Mendl, M. (2004). Animal behaviour—Cognitive bias and affective state. *Nature, 427*, 312.

Hayes, S., Hirsch, C.R., Krebs, G., & Mathews, A. (2010). The effects of modifying interpretation bias on worry in generalized anxiety disorder. *Behaviour Research and Therapy, 48*, 171–178.

Hertel, P.T., Brozovich, F., Joormann, J., & Gotlib, I.H. (2008). Biases in interpretation and memory in generalized social phobia. *Journal of Abnormal Psychology, 117*, 278–288.

Hertel, P.T., & El-Messidi, L. (2006). Am I blue? Depressed mood and the consequences of self focus for the interpretation and recall of ambiguous words. *Behavior Therapy, 37*, 259–268.

Hertel, P.T., Mathews, A., Peterson, S., & Kintner, K. (2003). Transfer of training emotionally biased interpretations. *Applied Cognitive Psychology, 17*, 775–784.

Hertel, P.T., Vasquez, E., Benbow, A., & Hughes, M. (in press). Recollection is impaired by the modification of interpretation bias. *Journal of Abnormal Psychology*.

Hirsch, C.R., Clark, D., & Mathews, A. (2006). Imagery and interpretations in social phobia: Support for the combined cognitive biases hypothesis. *Behavior Therapy, 37*, 223–236.

Hirsch, C.R., Hayes, S., & Mathews, A. (2009). Looking on the bright side: Accessing benign meanings reduces worry. *Journal of Abnormal Psychology, 118*, 44–54.

Hirsch, C.R., MacLeod, C., Mathews, A., Sandher, O., Siyani, A., & Hayes, S. (2011). The contribution of attentional bias to worry: Distinguishing the roles of selective engagement and disengagement. *Journal of Anxiety Disorders, 25*, 272–277.

Hirsch, C.R., & Mathews, A. (2000). Impaired positive inferential bias in social phobia. *Journal of Abnormal Psychology, 109*, 705–712.

Holmes, E.A., Lang, T.J., & Shah, D.M. (2009). Developing interpretation bias modification as a "cognitive vaccine" for depressed mood—Imagining positive events makes you feel better than thinking about them verbally. *Journal of Abnormal Psychology, 118*, 76–88.

Hoppitt, L., Mathews, A., Yiend, J., & Mackintosh, B. (2010a). Cognitive bias modification: The critical role of active training in modifying emotional responses. *Behaviour Therapy, 41*, 73–81.

Hoppitt, L., Mathews, A., Yiend, J., & Mackintosh, B. (2010b). Cognitive mechanisms underlying the emotional effects of bias modification. *Applied Cognitive Psychology, 24*, 312–325.

Isaacowitz, D.M., & Choi, Y. (2011). The malleability of age-related positive gaze preferences: Training to change gaze and mood. *Emotion, 11*, 90–100.

Jacoby, L.L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513–541.

Joormann, J., & Gotlib, I.H. (2007). Selective attention to emotional faces following recovery from depression. *Journal of Abnormal Psychology, 116*, 80–85.

Joormann, J., Hertel, P.T., LeMoult, J., & Gotlib, I.H. (2009). Training forgetting of negative material in depression. *Journal of Abnormal Psychology, 118*, 34–43.

Joormann, J., & Siemer, M. (2004). Memory accessibility, mood regulation, and dysphoria: Difficulties in repairing sad mood with happy memories? *Journal of Abnormal Psychology, 113*, 179–188.

Kornell, N., & Bjork, R.A. (2008). Learning concepts and categories: Is spacing the "enemy of induction"? *Psychological Science, 19*, 585–592.

Koster, E.H.W., De Raedt, R., Leyman, L., & De Lissnyder, E. (2010). Mood-congruent attention and memory bias in dysphoria: Exploring the coherence among information-processing biases. *Behaviour Research and Therapy, 48*, 219–225.

Koster, E.H.W., Fox, E., & MacLeod, C. (2009). Introduction to the special section on cognitive bias modification in emotional disorders. *Journal of Abnormal Psychology, 118*, 1–4.

Krebs, G., Hirsch, C.R., & Mathews, A. (2010). The effect of attention modification with explicit vs. minimal instructions on worry. *Behaviour Research and Therapy, 48*, 251–256.

Lawson, C., & MacLeod, C. (1999). Depression and the interpretation of ambiguity. *Behaviour Research and Therapy, 37*, 463–474.

Lawson, C., MacLeod, C., & Hammond, G. (2002). Interpretation revealed in the blink of an eye: Depressive bias in the resolution of ambiguity. *Journal of Abnormal Psychology, 111*, 321–328.

LeMoult, J., Hertel, P.T., & Joormann, J. (2010). Training the forgetting of negative material: The role of active suppression and the relation to stress reactivity. *Applied Cognitive Psychology, 24*, 365–375.

Lyubomirsky, S., Caldwell, N.D., & Nolen-Hoeksema, S. (1998). Effects of ruminative and distracting responses to depressed mood on retrieval of autobiographical memories. *Journal of Personality and Social Psychology, 75*, 166–177.

Mackintosh, B., Mathews, A., Yiend, J., Ridgeway, V., & Cook, E. (2006). Induced biases in emotional interpretation influence stress vulnerability and endure despite changes in context. *Behavior Therapy, 37*, 209–222.

MacLean, K.A., Ferrer, E., Aichele, S.R., Bridwell, D.A., Zanesco, A.P., Jacobs, T.L., . . . Saron, C.D. (2010). Intensive meditation training improves perceptual discrimination and sustained attention. *Psychological Science, 21*, 829–839.

MacLeod, C., & Cohen, I.L. (1993). Anxiety and interpretation of ambiguity: A test comprehension study. *Journal of Abnormal Psychology, 102*, 238–247.

MacLeod, C., Koster, E.H., & Fox, E. (2009). Whither cognitive bias modification research? Commentary on the special section articles. *Journal of Abnormal Psychology, 118*, 89–99.

MacLeod, C., & Mathews, A. (2004). Selective memory effects in anxiety disorders: An overview of research findings and their implications. In D. Reisberg & P. Hertel (Eds.), *Memory and emotion* (pp. 155–185). New York, NY: Oxford University Press.

MacLeod, C., Mathews, A., & Tata, C. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology, 95*, 15–20.

MacLeod, C., Rutherford, E., Campbell, L., Ebsworth, G., & Holker, L. (2002). Selective attention and emotional vulnerability: Assessing the causal basis of their association through the experimental manipulation of attentional bias. *Journal of Abnormal Psychology, 111*, 107–123.

Mathews, A., Fox, E., Yiend, J., & Calder, A. (2003). The face of fear: Effects of eye gaze and emotion on visual attention. *Visual Cognition, 10*, 823–836.

Mathews, A., & Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *Journal of Abnormal Psychology, 109*, 602–615.

Mathews, A., & MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annual Review of Clinical Psychology, 1*, 167–195.

Matt, G.E., Vazquez, C., & Campbell, W.K. (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review, 12*, 227–255.

McDaniel, M.A., & Callender, A.A. (2010). Cognition, memory, and education. In H.L. Roediger, III (Ed.), *Cognitive psychology of memory* (pp. 819–844). Oxford, England: Elsevier.

Mendl, M., Brooks, J., Basse, C., Burman, O., Paul, E., Blackwell, E., & Casey, R. (2010). Dogs showing separation-related behaviour exhibit a "pessimistic" cognitive bias. *Current Biology, 20*, 839–840.

Mineka, S., & Zinbarg, R. (2006). A contemporary learning theory perspective on the etiology of anxiety disorders: It's not what you thought it was. *American Psychologist, 61*, 10–26.

Mitte, K. (2008). Memory bias for threatening information in anxiety and anxiety disorders: A meta-analytic review. *Psychological Bulletin, 134*, 886–911.

Mogg, K., Bradbury, K.E., & Bradley, B.P. (2006). Interpretation of ambiguous information in clinical depression. *Behaviour Research and Therapy, 44*, 1411–1419.

Mogg, K., Millar, N., & Bradley, B.P. (2000). Biases in eye movements to threatening facial expressions in generalized anxiety disorder and depressive disorder. *Journal of Abnormal Psychology, 109*, 695–704.

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132*, 297–326.

Morris, C.D., Bransford, J.D., & Franks, J.J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*, 519–533.

Mystkowski, J.L., Craske, M.G., Echiverri, A.M., & Labus, J.S. (2006). Mental reinstatement of context and return of fear in spider-fearful participants. *Behavior Therapy, 37*, 49–60.

Needham, D.R., & Begg, I.M. (1991). Problem-oriented training promotes spontaneous analogical transfer: Memory-oriented training promotes memory for training. *Memory & Cognition, 19*, 543–557.

Nokes, T.J. (2009). Mechanisms of knowledge transfer. *Thinking & Reasoning, 15*, 1–36.

Nolen-Hoeksema, S., & Morrow, J. (1993). The effect of rumination and distraction on naturally occurring depressed moods. *Cognition & Emotion, 7*, 561–570.

Nolen-Hoeksema, S., Wisco, B.E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science, 3*, 400–424.

Parrott, W.G., & Hertel, P.T. (1999). Research methods in cognition and emotion. In T. Dalgleish & M. Power (Eds.), *The handbook of cognition and emotion* (pp. 61–81). Chichester, England: Wiley.

Peters, K.D., Constans, J.I., & Mathews, A. (in press). Experimental modification of attribution processes. *Journal of Abnormal Psychology*.

Richards, A., & French, C.C. (1992). An anxiety-related bias in semantic activation when processing threat/neutral homographs. *The Quarterly Journal of Experimental Psychology, 45A*, 503–525.

Rinck, M., Kwakkenbos, L., Dotsch, R., Wigboldus, D.H.J., & Becker, E.S. (2010). Attentional and behavioural responses of spider fearfuls to virtual spiders. *Cognition & Emotion, 24*, 1199–1206.

Roediger, H.L., III., & Karpicke, J.D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*, 181–210.

Roseman, I.J. (1991). Appraisal determinants of discrete emotions. *Cognition & Emotion, 5*, 161–200.

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

Russo, R., Whittuck, D., Roberson, D., Dutton, K., Georgiou, G., & Fox, E. (2006). Mood-congruent free recall bias in anxious individuals is not a consequence of response bias. *Memory, 14*, 393–399.

Salemink, E., Hertel, P., & Mackintosh, B. (2010). Interpretation training influences memory for prior interpretations. *Emotion, 10*, 903–907.

Salemink, E., & van den Hout, M. (2010). Trained interpretative bias survives mood change. *Journal of Behavior Therapy and Experimental Psychiatry, 4*, 310–315.

Salemink, E., van den Hout, M., & Kindt, M. (2007). Trained interpretive bias and anxiety. *Behavior Research and Therapy, 45*, 329–340.

Salemink, E., van den Hout, M., & Kindt, M. (2010). Generalisation of modified interpretive bias across tasks and domains. *Cognition & Emotion, 24*, 453–464.

Schartau, P.E.S., Dalgleish, T., & Dunn, B.D. (2009). Seeing the bigger picture: Training in perspective broadening reduces self-reported affect and psychophyiological response to distressing films and autobiographical memories. *Journal of Abnormal Psychology, 118*, 15–27.

Schmidt, N.B., Richey, J.A., Buckner, J.D., & Timpano, K.R. (2009). Attention training for generalized social anxiety disorder. *Journal of Abnormal Psychology, 118*, 5–14.

See, J., MacLeod, C., & Bridle, R. (2009). The reduction of anxiety vulnerability through the modification of attentional bias: A real-world study using a home-based cognitive bias modification procedure. *Journal of Abnormal Psychology, 118*, 65–75.

Smith, C.A., & Lazarus, R.S. (1993). Appraisal components, core relational themes, and the emotions. *Cognition & Emotion, 7*, 233–269.

Smith, C.A., & Pope, L.K. (1992). Appraisal and emotion: The interactional contributions of dispositional and situational factors. In M.S. Clark (Ed.), *Emotion and social behaviour* (pp. 32–62). Thousand Oaks, CA: SAGE.

Standage, H., Ashwin, C., & Fox, E. (2010). Is manipulation of mood a critical component of cognitive bias modification procedures? *Behaviour Research and Therapy, 48*, 4–10.

Teachman, B.A., & Addison, L.M. (2008). Training non-threatening interpretations in spider fear. *Cognitive Therapy and Research, 32*, 448–459.

Tran, T., Hertel, P.T., & Joormann, J. (2011). Cognitive bias modification: Induced interpretive biases affect memory. *Emotion, 11*, 145–152.

Vassilopoulos, S., Banerjee, R., & Prantzalou, C. (2009). Experimental modification of interpretation bias in socially anxious children: Changes in interpretation, anticipated interpersonal anxiety, and social anxiety symptoms. *Behaviour Research and Therapy, 47*, 1085–1089.

Velten, E. (1968). A laboratory task for induction of mood states. *Behaviour Research and Therapy, 6*, 473–482.

Watkins, E.R., Baeyens, C.B., & Read, R. (2009). Concreteness training reduces dysphoria: Proof -of-principle for repeated

cognitive bias modification in depression. *Journal of Abnormal Psychology, 118*, 55–64.

Wells, T.T., Beevers, C.G., Robison, A.E., & Ellis, A.J. (2010). Gaze behavior predicts memory bias for angry facial expressions in stable dysphoria. *Emotion, 10*, 894–902.

Wiers, R.W., Eberl, C., Rinck, M., Becker, E.S., & Lindenmeyer, J. (2011). Retraining automatic action tendencies changes alcoholic patients approach bias for alcohol and improves treatment outcome. *Psychological Science, 22*, 490–497.

Williams, J.M.G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin, 120*, 3–24.

Wilson, E.J., MacLeod, C., Mathews, A., & Rutherford, E.M. (2006). The causal role of interpretive bias in anxiety reactivity. *Journal of Abnormal Psychology, 115*, 103–111.

Yiend, J., & Mathews, A. (2001). Anxiety and attention to threatening pictures. *The Quarterly Journal of Experimental Psychology, 54A*, 665–681.