Using the diffusion model to investigate the cognitive processes responsible for attentional and interpretation biases in anxiety and depression: The benefits outweigh the challenges

by

Amy Manuel

Thesis
Submitted to Flinders University
for the degree of

Doctor of Philosophy (Clinical Psychology)

College of Education, Psychology, & Social Work

9/5/2018

Contents

Summary

Attentional and interpretation biases have long been considered a contributing factor to anxiety and depression. As such, research has turned to investigating ways to modify these biases to alleviate the distressing symptoms of anxiety and depression. However, the cognitive processes that are responsible for attentional and interpretation biases, and their successful modification, are not well understood. Furthermore, there are questions surrounding the reliability of the attentional bias score. This score is the primary measure of attentional bias; thus, the issues of reliability are a concern for the integrity of the findings in the attentional bias modification literature.

Through a series of studies, this thesis explores the value of applying a mathematical model, the diffusion decision model (Ratcliff, 1978), to data returned from the dot probe and yes/no tasks, two measures of biased attention toward, and interpretation of, emotional information, respectively. In doing so, the aim is to advance theoretical understanding of the mechanisms that underlie cognitive biases in anxiety and depression. The diffusion decision model belongs to a class of models called evidence accumulation models. Evidence accumulation models are a more sophisticated form of analysis that can isolate and identify different implicit decisional processes. By doing so, better understanding of these processes can potentially lead to the development of more targeted interventions, and ultimately improve treatment outcomes for individuals living with anxiety and depression.

Additionally, this thesis explores the potential of the diffusion decision model to be a more sensitive and reliable way to measure attentional bias. If the suitability of diffusion decision model analysis as an alternative, more reliable, way to analyse data from the dot probe task is established, the findings presented in this thesis may encourage other researchers without a mathematical psychology background to explore the benefits of these kinds of models in their own work.

This thesis presents the successful application of diffusion decision model analysis to data from the dot probe task for the first time. In doing so, the capacity for the diffusion model to identify implicit decisional processes that differ between anxiety and depression that are not

captured by RTs alone, has been demonstrated. While the test-retest reliability of the diffusion model parameters was mixed, guidance has been provided for future research to gain clarity on the reliability of the diffusion model parameters, and their suitability as an alternative measure of attentional bias. Finally, the research in this thesis has demonstrated the value of adopting a mathematical psychology analytical approach to the field of applied clinical psychology.

DECLARATION

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Signed

Date: 15/10/2017

Acknowledgements

To my supervisors, Professor Eva Kemps, and Professor Jason McCarley.

Thank you for your support, wisdom, humour, and faith that I would get this done.

To my family, friends, and colleagues. Thank you.

To the staff at Flinders University who have had a part to play in the completion of this research.

Thank you.

To the participants of this research. Thank you.

Chapter 1: Introduction

Anxiety and depression have the highest prevalence rate of all mental health issues, affecting approximately 10% of the world's population (Chisholm et al., 2016). According the Diagnostic and Statistical Manual of mental disorders (5th Edn; American Psychiatric Association, 2013), generalised anxiety disorder is characterised by persistent, ongoing, excessive worry, that is difficult to control. Symptoms can include restlessness, fatigue, difficulty concentrating, irritability, and sleep disturbance, which cause significant distress in daily functioning. Correspondingly, major depressive disorder is characterised by uncharacteristic levels of sadness or hopelessness. Symptoms can include diminished interest in previously pleasurable activities, appetite change, weight gain or loss, psychomotor agitation or retardation, fatigue, feelings of worthlessness, indecision, poor concentration, recurrent thoughts of death, which cause significant distress in daily functioning.

According to Beck's Cognitive Theory of Emotional Disorders (1976), individuals are drawn to mood-congruent emotion; this is known as the *Content Specificity Hypothesis*. Specifically, individuals with anxiety are prone to attend to stimuli they perceive as threatening, and interpret benign information as more threatening, whereas individuals with depression are drawn to dysphoric stimuli, and interpret benign or ambiguous information with a dysphoric overtone. These biases in attention for mood-congruent stimuli, termed attentional and interpretation biases, have been hypothesised to have an aetiological and maintaining role in anxiety and depression. As such, research into the processes by which these biases operate, and how to effectively reduce their impact, has been expanding. One such area of research is attentional bias modification.

Attentional bias modification developed from attempts to induce an attentional bias to try to understand the causal role it has on mood (MacLeod, Mathews, & Tata, 1986). Over time, the research focus has shifted to how this bias induction could be used to effectively reduce, or neutralise, negative attentional biases. Attentional bias modification (ABM; MacLeod, Rutherford, Campbell, Ebsworthy, & Holker, 2002) is now being promoted as a potential therapeutic technique

that has the ability to train away negative bias, using a simple, implicit computer task called the dot probe task.

The dot probe task is a commonly used attentional bias assessment tool. It is a computerised task that instructs participants to focus on a fixation cross in the centre of a computer screen. The fixation interval is followed by the brief presentation of two stimuli, typically images or words, side-by-side or vertically aligned. One stimulus is emotional and the other is neutral in valence. After the stimuli disappear, a dot appears in the location previously occupied by one of the stimuli, and the participant is required to execute a speeded response indicating in which of the two positions the probe appeared. The participant is instructed to ignore the stimuli and focus only on the location of the probe. Trials on which the dot replaces the emotional stimulus are called congruent trials, and trials on which the dot replaces the neutral stimulus are called incongruent trials. There are equal numbers of congruent and incongruent trials for all emotional categories across the task. If the participant's attention has been captured by a specific emotional stimulus, it is assumed to be reflected by faster RTs when the dot appears in the location previously occupied by that stimulus. An estimate of attentional bias for emotional stimuli is calculated by subtracting the mean RT for congruent trials from the mean RT for incongruent trials. Positive scores result from faster RTs to congruent trials and are taken to indicate an attentional bias toward that emotion.

In the training paradigm that has been developed from this task, the probe appears most often at the location of the stimulus type that is the focus of training (e.g., neutral), and only infrequently at the location of the stimulus type that is the target of modification (e.g., negative). It is hypothesised that training effects occur on an implicit level, which has numerous positive implications for a clinical treatment tool. For example, one of the common symptoms of anxiety is hypervigilance. Hypervigilance reduces cognitive capacity for higher-order thinking. Similarly, a common symptom of depression is poor concentration, another higher-order cognitive process. Limitations on an individual's higher-order cognitive processes reduce their capacity to engage with forms of therapy that rely on these processes, e.g., cognitive behavioural therapy (Renaud, Russell,

& Myhr, 2014). This can hinder therapeutic outcomes, which can exacerbate symptoms further due to feelings of failure or incompetence. The availability of a treatment tool that could reduce bias and mitigate the symptoms that interfere with other forms of therapy would be incredibly valuable.

The promise of a treatment tool to target implicit attentional biases is clearly encouraging. However, the cognitive processes that are responsible for the occurrence of an attentional bias, or for successful bias modification, are still not fully understood. The attentional bias score can identify differences between groups, and whether a modification paradigm has shifted bias values because of training, but the processes that underlie the group differences or the change over time cannot be identified. Additionally, one of the major concerns regarding the dot probe task is that the reliability of the attentional bias score is poor. This has implications for the strength of the evidence for the efficacy of ABM, because confidence in research findings can only be as robust as the psychometric properties of the task used. Recommendations have been made to address the reliability of the attentional bias score and these include: (a) finding a new task, (b) being open to the possibility that the theory that attentional bias as a stable construct is flawed, and (c) applying a more sophisticated analysis technique that may have better sensitivity to reliably detect change (Price et al., 2015). It is this last suggestion that is addressed in this thesis.

Evidence accumulation models are a more sophisticated form of analysis that can isolate and identify different decisional processes, but they can be difficult to fit to data. As such, much of the work analysing cognitive processes in clinical disorders using evidence accumulation models has come from researchers with a strong mathematical psychology background (e.g. Heathcote et al., 2015; White, Ratcliff, Vasey, & McKoon, 2010; White, Skokin, Carlos, & Weaver, 2016). However, with several methods of fitting the models to data now available to researchers (Fast-dm, Voss & Voss, 2007; EZ diffusion, Wagenmakers, van der Maas, & Grasman, 2007; DMAT, Vandekerckhove & Tuerlinckx, 2008), the models are becoming easier to use and apply for researchers without a mathematical psychology background. The program *Fast-dm-30 (*Voss & Voss, 2015) was used to fit the data in the studies that form this thesis.

Aims

The aims of this thesis are threefold: (1) To successfully fit the diffusion model to data from the dot probe task; (2) To advance understanding of the cognitive processes responsible for attentional and interpretation biases in anxiety and depression; and (3) To assess the reliability of the diffusion model parameters, and their suitability as a reliable alternative to the traditional dot probe measure.

Contribution of the Current Thesis

The primary contribution of this thesis lies in the consolidation of two areas of psychology, applied cognitive psychology and mathematical psychology, for the advancement of clinical psychological outcomes. Using an analytical method that is often shied away from in applied psychology due to its complexity, this thesis explores the value of the diffusion model to advance theoretical understanding of the mechanisms that underlie cognitive biases in anxiety and depression. By doing so, these processes can lead to the development of more targeted interventions, and ultimately improve treatment outcomes for individuals living with these disorders.

Due to the poor psychometric properties of the attentional bias score, an alternative, more reliable measure is needed. If the suitability of fast-dm analysis as an alternative, more reliable, way to analyse data from the dot probe task is established, the findings presented in this thesis may encourage other researchers without a mathematical psychology background to explore the benefits of these kinds of models in their own work.

Thesis Structure

The thesis is arranged over six chapters. Chapter 1 provides the context for the subsequent research chapters. It provides a broad overview of the value of attentional bias modification, as well as the methodological obstacles of the primary measure of attentional bias, the dot probe task. It also outlines the aims and the important contribution of this thesis. Chapter 2 presents an overview

of the current issues surrounding research into attentional bias and its modification, in anxiety and depression. Namely, how the poor reliability of the attentional bias score, derived from the dot probe task, impacts the integrity of research in the field, and the inability of the bias score to offer insight into the cognitive processes that it is hypothesised to represent. From this, an argument is presented as to why a novel form of analysis in this area, the diffusion decision model, may be able to address these issues. The next three chapters present separate studies prepared for publication. Chapter 3 presents the first research study. The diffusion decision model, using the fast-dm 30 program, is applied to data from the dot probe task in a non-clinical population to assess its suitability to account for the data from the dot probe task. This is the first time that data from the dot probe task have been analysed using the diffusion model. The results were promising and encouraged further application of the model. Chapter 4 (Study 2) broadens the line of enquiry to interpretation bias. The diffusion model is applied to data from a yes/no emotional interpretation bias task to identify the cognitive processes that are responsible for the interpretation of ambiguous stimuli in a non-clinical population. The model was applied successfully and offered further insight into the processes underlying cognitive biases. Chapter 5 presents a bias modification study. The aim was to examine the effects of bias modification on the fast-dm parameters returned for attentional and interpretation bias, in individuals with a history of anxiety and depression. Changes in parameters from pre- to post-training would identify which cognitive processes underlie these cognitive biases. Unfortunately, neither a disorder-congruent attentional nor interpretation bias was identified from the data obtained at the pre-training testing phase, so this was not achieved. What was achieved was an examination of the cognitive processes that differ across the attentional and interpretation bias tasks, between clinical and non-clinical cohorts. This study also evaluates the test-retest reliability of the diffusion model parameters from the data from a control cohort taken at pre- at post-training. Finally, Chapter 6 draws together the pertinent findings from the three studies to respond to the aims of this thesis. The theoretical and clinical implications of the findings are

explored, as are the challenges that arise when trying to bring a more complex model of analysis into the broader research domain. From this, directions for future research are proposed.

Chapter 2: Literature Review

Measurement and analysis of attentional bias in clinical research on anxiety and depression: why evidence accumulation models may be the way forward.

Amy Manuel, Eva Kemps and Jason S. McCarley

School of Psychology, Flinders University, Adelaide, Australia

Correspondence concerning this article should be addressed to Amy Manuel, School of Psychology, Flinders University, GPO Box 2100, Adelaide, SA 5001, Australia. Electronic mail may be sent to Amy.Manuel@flinders.edu.au

Abstract

Anxiety and depression have the highest prevalence rate of all mental health issues, for which an attentional bias for mood-congruent stimuli has been implicated in the onset and maintenance of both disorders. Attentional bias modification (ABM) has been promoted as a technique designed to redress this negative bias, using a simple, implicit computer task called the dot probe task. The implications of this task as a treatment tool are valuable in that engagement with other, more explicit therapies can be one of the biggest challenges a therapist faces with these clients. However, the primary task used to measure the success of ABM, the dot probe task, has recently been awash with questions surrounding its reliability and the implications this has on the strength of evidence for the efficacy of ABM. We review the literature and identify inconsistent findings and methodological issues concerning this task, as well as questions surrounding the theoretical nature of the mechanisms of change in successful ABM. To address these issues, we recommend the use of evidence accumulation models to analyse the data returned from the dot probe task. Evidence accumulation models can isolate and identify different implicit decisional processes. We explain how these models may clarify the inconsistencies in findings of between-group differences, as well as provide information about what occurs at the level of the individual. Furthermore, we discuss the clinical implications of this more comprehensive method of analysis, and its potential to identify individuals for whom ABM may be a particularly effective intervention.

Key words: Anxiety, depression, attentional bias, attentional bias modification (ABM), dot probe task, evidence accumulator models

Measurement and analysis of attentional bias in clinical research on anxiety and depression: A way forward

Anxiety and depression affect nearly 10% of the world's population at any one time, costing the global economy an estimated $US1 trillion dollars each year (Chisholm et al., 2016). There is an estimated four-fold dollar return on treatment, such that for every dollar spent, an average of $4 is gained in health benefits and employment outcomes (Chisholm et al., 2016). Therefore, investing in effective treatment for anxiety and depression is beneficial not only for the individual, but also from a socio-economic perspective. NICE (2011) guidelines recommend cognitive behavioural therapy (CBT) as the gold standard treatment for anxiety and depressive disorders. However, even with successful rounds of treatment, both disorders have high rates of relapse (Burcusa & Iacono, 2007; Scholten et al., 2013). As a result, researchers continuously seek to improve treatment efficacy and longevity, by exploring different treatment types and delivery modes, or targeting different underlying mechanisms. CBT specifically targets explicit cognitive processes to alleviate symptoms of anxiety and depression; however, it has long been theorized that implicit cognitive processes are involved in the onset and maintenance of both anxiety and depression (e.g., Beck, Rush, Shaw, & Emery, 1979; Beck, Emery, & Greenberg, 1985). For this reason, research has begun to focus on therapeutic techniques that target implicit cognitive processes in conjunction with CBT, with the aim of improving treatment outcomes. One such technique is *attentional bias modification (ABM)*. The clinical advantage of ABM is that it is hypothesised to work at the implicit cognitive level, and thus, the brevity and computerised, automatic nature of the training, means the task is not cognitively onerous, unlike other forms of therapy, and there is the potential for clients to complete the training daily in their own surroundings, with the further benefit of increased client engagement.

Attentional bias is the tendency to allocate attention preferentially toward certain kinds of information in the environment at the expense of other information (MacLeod, Mathews, & Tata, 1986). The dot probe is the task most commonly used to measure attentional bias, and is the task that was used to establish a link between attentional bias and clinical disorders such as anxiety and

depression (Bradley, Mogg, & Lee, 1997; MacLeod, Rutherford, Campbell, Ebsworthy, & Holker, 2002; Mobini, Reynolds, & Mackintosh, 2012; Mogg & Bradley, 2005). The evidence of an association between attentional bias and symptoms of anxiety and depression led to research on ABM, which attempts to reduce clinical symptomology by reducing an individual's automatic preferential attention to negative, disorder-relevant stimuli.

In recent years, there have been numerous reports of ABM treatment success (e.g. MacLeod & Clarke, 2015). The cognitive processes that underlie changes in attentional bias scores are hypothesised to stem from a shift in implicit bias; however, current methods of analysis cannot reliably confirm this. More worryingly, a recent meta-analysis by Cristea, Kok, and Cuijpers (2015) posits that claims of successful ABM have been made too soon, and are based on substandard research methodology. For example, the reliability of the dot probe task has been questioned (Schmukle, 2005; Staugaard, 2009), which, ultimately casts doubt over the validity of the measure and the results inferred from it. In addition, there are concerns that false positive findings and a file drawer effect may have resulted in a biased representation of the efficacy of ABM, such that the effect seems stronger and more robust than it may be.

This paper reviews the literature and identifies inconsistent findings and methodological issues concerning to the use of the dot probe task to measure attentional bias. We also consider the theoretical nature of the mechanisms of change in successful ABM, and, following on from a suggestion by Price and colleagues (2015), recommend the use of a more informative means of analysing the data obtained from the dot probe task. Specifically, we advocate for the use of evidence accumulation models to analyse behavioural data. In doing so, we outline how evidence accumulation models may reveal the cognitive mechanisms of attentional bias and successful bias modification, and we describe the clinical implications of this knowledge. Further, we explain how evidence accumulation models have the potential to clarify the inconsistencies in findings of between-group differences, and we offer insight into what occurs at the level of the individual. Individual difference data are of particular interest because this is where the reliability of the dot

probe task is most strongly questioned. These are also the data that will inform clinicians' use of ABM as a treatment for their individual clients.

Attentional Bias

Theories of selective attention refer to the cognitive mechanisms that select and prioritize information for processing, enabling us to notice, recognise, or act on some elements of the environment to the exclusion of others (Broadbent, 1958; Treisman, 1960; Deutsch and Deutsch, 1963). When functioning well, selective attention guides our limited perceptual and cognitive resources toward task-relevant information to ensure adaptive behaviour. Sometimes, however, this mechanism may become overly attuned to specific stimuli, exacerbating associated emotional responses (Beck, 1976; Mathews & MacLeod, 2005). Accordingly, when the stimulus that is excessively attuned to is negative in nature, the individual is said to have an attentional bias for negative information (Fox, Russo, & Georgiou, 2005; MacLeod, Campbell, Rutherford, & Wilson, 2004; Mathews & MacLeod, 2005). This bias is an automatic process that is neither within conscious awareness nor under conscious control (Beard, 2011).

An attentional bias for negative information has been implicated in the aetiology and maintenance of both anxiety and depression (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & Van Ijzendoorn, 2007; Beck, 1976; Mathews & Mackintosh, 1998; Mathews & MacLeod, 2005; Mogg & Bradley, 2005; Pergamin-Hight, Naim, Bakermans-Kranenburg, Van Ijzendoorn, & Bar-Haim, 2015), but presents differently for each disorder (Browning, Holmes, & Harmer, 2010). Anxiety has been associated with a bias toward threatening stimuli whereas depression has been associated with a bias toward dysphoric stimuli (Mogg & Bradley, 2005). Additionally, the stage of attentional processing at which the bias appears to operate differs between the two disorders (Mogg & Bradley, 2005). Specifically, individuals with a variety of anxiety disorders have a stronger orientation toward stimuli that are perceived as threatening than do individuals who do not have anxiety. That is, when presented with stimuli thought to be threatening in nature, individuals with anxiety will have their attention captured more quickly by those stimuli than their non-anxious

counterparts (Koster, Verschuere, Crombez, & Van Damme, 2005; Pergamin-Hight et al., 2015; Shechner et al., 2012). Individuals with depression do not always attend to dysphoric stimuli at a faster rate than their non-depressed counterparts, but they do have greater difficulty disengaging their attention from the dysphoric stimulus once it has captured their attention (Bradley et al., 1997; Donaldson, Lam, & Mathews, 2007; Gotlib, Krasnoperova, Yue, & Joormann, 2004; Leyman, De Raedt, Schacht, & Koster, 2007).

The attentional biases for negative information found in anxiety and depression are evident in different windows of time following stimulus presentation: anxiety is associated with an attentional bias for negative information when the probe is presented 10 to 500ms after stimulus onset, whereas for depression the bias is most commonly identified when the probe appears between 500 and 1,000ms after stimulus onset (Gotlib et al., 2004). This stimulus onset asynchrony (SOA) between the attention capturing stimulus and the probe accounts for when (deployment versus disengagement) the bias is found, and has been explained by dual process models of attention. These models propose two systems of attention: implicit, automatic, "bottom-up" processing, and explicit, controlled, "top down" processing, which are temporally separable (Mathews & MacLeod, 2005; & see Carver, Johnson, & Joormann, 2008, for a review of dual process models in emotional disorders). The heightened initial engagement of attention in anxiety, and the increased difficulties in disengaging attention in depression, suggest that anxiety is more strongly associated with abnormalities in the early, automatic processing of information, whereas depression is more strongly associated with abnormalities in the later, controlled, processing stages (De Raedt & Koster, 2010; Leyman et al., 2007).

Measurement of Attentional Bias

The dot probe task is the most commonly used method for measuring attentional bias. The task was adapted by MacLeod and colleagues (1986) from experimental cognitive psychology paradigms such as the Posner task (Posner, 1980), which enables the orientation of visuospatial attention to be assessed from manual response times (RTs) to visual probes. The premise is that

individuals respond faster to a probe stimulus (e.g., a small dot) that is presented in an attended rather than unattended region of a display.

Since its initial development, there have been numerous modifications to the dot probe task (Koster, Crombez, Verschuere, & De Houwer, 2004; Mogg, Bradley, De Bono, & Painter, 1997), and as a result the task exists in several variations. Most commonly, a trial begins with a fixation cross in the centre of the computer screen. Two stimuli are then presented simultaneously, equidistant from the centre of the screen. The stimuli are either words or pictures associated with the construct of interest (e.g., in the case of anxiety, threatening information, and in the case of depression, dysphoric information) or words or pictures that are considered benign (neutral stimuli). Stimulus pairs generally consist of one of each of a negative and neutral item. The stimuli remain onscreen briefly (duration across studies has ranged from 100 – 1500ms; the most common presentation is 500ms), and when they disappear, a probe (e.g., dot, arrow, letter) appears in the location previously occupied by one of the stimuli. Generally, the probe is equally likely to replace either the negative or neutral stimulus. The participant is then required to make a forced-choice response, indicating either the position or the identity of the probe.

The bias score is calculated by subtracting the mean RT to probes replacing the emotional target (i.e., congruent trial) from the mean response time to probes replacing the neutral target (i.e., incongruent trial). Positive bias scores reflect shorter RTs for probes that replace emotional stimuli, and negative bias scores reflect shorter RT for probes that replace neutral stimuli. Positive scores imply that attention was drawn toward the emotional stimulus, whereas negative scores imply an orienting away from the emotional stimulus, or attentional avoidance. When an individual has a positive score, they are said to have an attentional bias to the stimuli of interest, in this case, information that is negative in valence.

Variations to the dot probe protocol can include differences in stimulus type (e.g., lexical versus visual), stimulus valence (e.g., positive/negative), emotional intensity of the stimulus, and SOA (under the assumption that a shorter SOA [100-500ms] targets earlier, automatic stages of

processing, identifying attentional capture, and that a longer SOA [>1000ms] targets later, more strategic stages of processing, identifying difficulty disengaging attention). Variations in participant characteristics can include differences in disorder type (e.g., anxiety versus depression) and sample characteristics (e.g., clinical versus subclinical versus nonclinical). Characteristics of the dot probe protocol and the participant sample are important moderating variables that make between-study comparisons difficult (Cisler, Bacon, & Williams, 2009). As a result, the reliability of the evidence obtained by the dot probe task continues to be debated, as does the temporal stage of attentional bias activation, and the best way to measure and analyse the behavioural data.

Attentional Bias Modification

To understand how attentional bias contributes to the aetiology and maintenance of anxiety and depression, researchers began to examine the effect of inducing an attentional bias for negative information and modifying existing negative attentional biases on an individual's emotional wellbeing. Bias modification research was initially used to demonstrate the causal nature of the relationship between attentional bias and emotional stimuli (MacLeod et al., 2002). By training participants to attend toward a negative emotional stimulus, thus inducing an attentional bias for negative information, the effects of their responses to emotional stressors increased and the causal nature attentional bias on emotional reactivity was inferred (MacLeod et al., 2002). Focus then turned to the ability of ABM to alleviate clinical symptomology. By training participants to attend to an alternative, neutral stimulus, the aim is to inhibit attentional capture by the negative stimulus toward which their attention is biased. This ABM training aims to reduce the associated emotional response that the stimulus triggers. As time and research have evolved, the potential implication of bias malleability for clinical populations has grown into what is now known as ABM therapy.

The most commonly used task to modify attentional bias is a variant of the dot probe task that was devised by MacLeod and colleagues (2002). It is hypothesized to directly alter the implicit cognitive processes that result in an attentional bias to disorder-congruent stimuli. In the assessment version of the dot probe task, probes are presented equally often following each stimulus type

(positive/neutral/negative). In the training variant, the probe appears most often at the location of the stimulus type that is the focus of training (e.g., neutral), and only infrequently at the location of the stimulus type that is the target of modification (e.g., negative). Results of bias modification, and its subsequent effect on clinical symptoms, have been mixed (Clarke, Notebaert, & MacLeod, 2014; Cristea et al., 2015; Kuckertz and Amir, 2015; MacLeod and Clarke, 2015). This is perhaps not surprising given the debate surrounding the ability of the dot probe task to capture bias, and the hypothesized function of attentional bias within clinical disorders. While there is more evidence for attentional bias in anxiety than in depression (Mogg & Bradley, 2005), the underlying issues are the same: the mechanisms of how the bias operates and its relationship to disorder symptomology are still not fully understood.

Progression of Attentional Bias Research over the Last Decade

Attentional bias for threat, as measured by the dot probe paradigm, has been reported in individuals with anxiety since the 1980s (e.g., MacLeod et al., 1986), and attentional bias for dysphoric information in those with depression since the 1990s (e.g., Mogg, Bradley, & Williams, 1995). Early this century, efforts to modify attentional bias and explore its causal effect on symptoms of anxiety and depression began (e.g., MacLeod et al., 2002). As this early research into the efficacy of ABM returned favourable findings (e.g., Schmidt, Richey, Buckner, & Timpano, 2009; See, MacLeod, & Bridle, 2009; Watkins, Baeyens, & Read, 2009) and a special section in a journal (Goodman, 2009) fostered excitement, research into ABM exploded. This research boom produced enough data for narrative reviews, and eventually, for quantitative meta-analytic reviews. Unfortunately, these reviews have discovered that the efficacy of ABM may be less robust than initial excitement indicated.

Overview of the evidence of attentional bias in anxiety and depression. Over the last decade, numerous narrative review papers and two quantitative meta-analyses have explored the literature on attentional bias in anxiety (Bantin, Stevens, Gerlach, & Hermann, 2016; Bar-Haim et al., 2007; Cisler et al., 2009; Cisler & Koster, 2010; Mathews & MacLeod, 2005; Mogg & Bradley,

2005; Pergamin-Hight et al., 2015; Shechner et al., 2012; Van Bockstaele et al., 2014) and depression (Everaert, Koster, & Derakshan, 2012; Gotlib & Joormann, 2010; Leppänen, 2006; Mathews & MacLeod, 2005; Mogg & Bradley, 2005; Peckham, McHugh, & Otto, 2010). The consensus was that attentional biases exist for both anxiety and depression, but that the mechanisms by which they operate differ.

A meta-analysis (Bar-Haim et al., 2007) of 172 studies (N = 2,263 anxious, N = 1,768 non-anxious) concluded that attentional bias was reliably demonstrated in anxious populations with a moderate effect size of $d = 0.45$, but was not demonstrated in non-anxious populations. It also reported that the publication bias was not a concern as the number of studies included in the meta-analysis was 20 times that of Rosenthal's (1991) fail safe. A more recent meta-analysis (Peckham et al., 2010) of 29 studies (N = 2,351, 1,459 of whom completed the dot probe task, and 892 of whom completed the emotional Stroop task) examined differences in attentional bias between dysphoric/depressed and non-depressed/non-dysphoric individuals. It found a moderate effect of attentional bias ($d = 0.52$), as measured by the dot probe task, to negative information in depressed relative to non-depressed individuals.

Different attentional bias mechanisms in anxiety and depression. Reviews by Mathews and MacLeod (2005) and Mogg and Bradley (2005) highlighted differences in the way attentional bias seemed to manifest in anxiety and depression. Mathews and MacLeod concluded that attentional bias operates rapidly and automatically in anxiety, emerging at SOAs of less than 500ms, but is slower and more controlled in depression, emerging at SOAs of greater than 500ms. Mathews and MacLeod posit the reason for this is that within anxiety, individuals maintain an attentional set for potential threat and engage with it quickly. However, in depression the sensitivity to dysphoric cues is inhibited, and thus a slower attentional processing strategically directed at mood congruent stimuli is engaged (i.e., a rumanitive response; Mathews & MacLeod, 2005).

Mogg and Bradley (2005) came to similar conclusions, noting that there is evidence of a non-conscious attentional bias toward negative stimuli in anxiety, but that the evidence of a similar

rapid, implicit attentional capture in depression is less convincing. They reported that when an attentional bias for negative information is found in individuals with depression, it is found at the later stages of processing that allow individuals to elaborate on self-relevant negative stimuli. As a result, Mogg and Bradley hypothesised that the mechanisms underlying attentional bias in anxiety and depression may differ, with depression related to a difficulty disengaging from self-relevant dysphoric information, as opposed to an attentional capture by threatening information, as is the case in anxiety.

However, as research into the role of attentional bias in anxiety and depression progressed, findings became more mixed. In a review of the literature, Cisler and colleagues (2009) noted that although anxious populations showed a bias toward threat early in stimulus processing, there was evidence that they also seemed to have difficulty disengaging from the threat and would avoid the threatening stimulus as processing continued (e.g., Koster et al. 2006; Salemink et al. 2007). As a result, Cisler and colleagues recommended that current models of information processing and attentional bias be adjusted to incorporate a temporal component. While facilitated attention was considered to be automatic, and difficulty disengaging was considered to be strategic (as per Garner, Mogg, & Bradley, 2006; Koster et al., 2004; Koster et al., 2005; Koster, Crombez, Verschuere, Van Damme, & Wiersema, 2006; Mogg et al., 1997; Mogg, Bradley, Miles, & Dixon, 2004), Cisler and colleagues proposed that attentional avoidance of threatening stimuli could be accounted for by Mogg, Mathews, & Weinman's (1987) *vigilance-avoidance* hypothesis. According to this hypothesis, attention initially orients to threat, then shifts away to prompt escape from the threatening stimulus or to reduce anxious or fearful affect. Thus, they concluded that differences in findings regarding attentional bias within the literature may reflect a measurement issue related to the experimental stimulus presentation duration, rather than evidence that the bias does not exist.

Cisler and Koster (2010) tried to make sense of this time-course complexity of attentional bias by relating the hypothesised cognitive and affective processes to neural mechanisms.

Specifically, they linked facilitated attention with threat detection, the amygdala, and automatic (bottom-up) processing. In addition, they linked attentional avoidance with emotion regulation goals, the prefrontal cortex, and strategic (top-down) processing. They further linked difficulty disengaging from threat to a lack of attentional control, and thus theorised that it sits somewhere between automatic and strategic processing, in that the activation of the threat detection system impeded the efficacy of the prefrontal cortex to down regulate. Although this hypothesised link to neural structures could account for the results seen in anxiety, it does not explain the rather different pattern of results for depression.

Gotlib and Joormann (2010) reported that the facilitated attention toward disorder congruent information found in individuals with anxiety appeared to be lacking in individuals with depression. However, once negative information had become the focus of their attention, depressed individuals had increased difficulty disengaging from it, and seemed to engage with it in a ruminative fashion. By contrast, Peckham and colleagues (2010) examined attentional bias differences between dysphoric/depressed and non-depressed/non-dysphoric individuals, and did not find evidence of a time course effect on bias, concluding that both early and late stage processing contribute to attentional bias in depression. De Raedt and Koster (2010) proposed a neurobiological framework to account for the mood-congruent attentional bias found at the later stages of information processing that characterizes depression. They suggested that difficulties in frontal lobe activation led to the amygdala being activated for longer periods of time, resulting in difficulty disengaging from negative information. Thus, while there have been significant developments in our hypotheses of the cognitive mechanisms that underlie attentional bias in anxiety and depression, there is still much that we do not know and our understanding remains at the theoretical level.

Progression of ABM Research over the Last Decade

Research into ABM has gained enough traction over the last decade that the first reviews of the efficacy of the technique have started to emerge. In the first such review, Bar-Haim (2010) concluded that ABM is successful at directly reducing attentional bias for threatening information,

and subsequently, associated symptoms of anxiety. He did, however, caution that the stability and test-retest reliability of the dot probe task had to be carefully established before treatment efficacy could be assured. That same year, the first quantitative meta-analysis into the effect of attentional bias modification on anxiety (Hakamata et al., 2010) reported that ABM training reduced anxiety significantly more than control training, with a medium effect size, d = .61. Hakamata and colleagues (2010) concluded that ABM was a promising treatment for anxiety, and they recommended that additional randomised controlled trials be carried out to evaluate the treatment efficacy in clinical populations. However, questions as to the proposed mechanisms responsible for the changes in bias scores and/or associated clinical symptomology began to emerge.

Mechanisms of change. When creating the modified version of the dot probe paradigm used to retrain attentional bias, MacLeod and colleagues (2002) initially hypothesized that successful bias modification stemmed from the implicit training of bottom-up attentional processes. However, in a review of the evidence of bias modification in anxiety, Beard (2011) noted that the effects of ABM do not always show on measures of anxiety or mood alone, but require a stressor to be evident. She argued that this indicated that ABM alters cognitive vulnerability to stress rather than directly affecting anxiety symptomology or mood. Beard suggested that while cognitive biases play a causal role in anxiety vulnerability, individual differences in the ability to control attention (top-down processing) are the key moderator of the relationship between attentional bias and anxiety in that anxious individuals with good attentional control do not exhibit an attentional bias to negative information. As such, she argued that ABM may improve attentional control overall (top-down), rather than targeting the stimulus-driven (bottom-up) components of attentional bias toward negative information. While this conclusion is contrary to the mechanisms of change initially hypothesised by MacLeod and colleagues (2002), it too presupposes a dual-process model of biased processing. Regardless of the mechanism of change underlying successful bias modification, ABM interventions that alter the function of either one or both systems (top-down/ bottom-up) will

influence attention to emotional stimuli and, in turn, current symptomatology of anxiety and depression.

Efficacy of ABM on anxiety and depression. Reviews continued to praise ABM as a potentially effective treatment for anxiety and depression (e.g., Baert, Koster, & De Raedt, 2011; Hertel & Mathews, 2011; MacLeod, 2012; MacLeod & Mathews, 2012), with a meta-analysis (Beard, Sawyer, & Hofmann, 2012) concluding that ABM has a moderate and robust effect on attentional bias toward threat stimuli. However, other reviews were published that raised some important questions regarding the mechanisms of action responsible for the successful results seen thus far, highlighting limitations of the research. For example, a meta-analysis investigating the efficacy of ABM for anxiety and depression noted that the small effect sizes observed (ranging from g =.13 to .29, k=15[1]) suggested that the efficacy of ABM may be more modest than had initially been believed (Hallion & Ruscio, 2011). This is potentially because most studies investigating treatment efficacy to date were pilot studies comprising small samples, resulting in unreliable effect sizes (Beard, 2011). Importantly, Beard also noted that only one published study (Amir et al., 2009) followed the Consolidated Standards of Reporting trials (CONSORT) guidelines. In addition, most studies used the dot probe task to both train the change in attentional bias and assess that change. As a result, a generalized change in bias could not be distinguished from a narrower, task-specific response.

Emmelkamp (2012) went so far as to call attentional bias modification 'the Emperor's new suit', arguing that regular cognitive therapy and exposure methods achieve much better clinical results for anxious individuals than attentional bias modification procedures. In line with this, Mogoaşe, David, and Koster (2014) conducted a meta-analysis (k = 43, N = 2,268) on the therapeutic benefit of ABM for anxiety, and ultimately concluded that the effect is, at best, small (g = 0.160), and that more efficient, psychometrically sound procedures are needed for assessing and modifying attentional bias than the dot probe task.

[1] g = Hedge's g. Effect sizes are comparable to Cohen's d. k = number of studies included in the meta-analysis

Cristea and colleagues (2015) conducted an updated meta-analysis examining the efficacy of cognitive bias modification (CBM) in general, so ABM and interpretation bias modification (IBM), across all samples (clinical and non-clinical) that have been included in CBM studies, as well as on clinical samples separately. Their conclusion of the state of the literature was unfavourable. They noted numerous weaknesses within the literature that limit the conclusions that could be made about the efficacy of CBM. Specifically, they reiterated Beard's (2011) concern that research in the field was overly reliant on small studies that did not follow the CONSORT guidelines. The results of their meta-analysis showed small, if any, effects of attentional bias modification (*k=10, g =.02, n.s.),* and they concluded that any CBM effects that were significant were because of outliers and/or publication bias and that adjustment for publication bias rendered any significant results unconvincing. Cristea and colleagues also reported evidence of demand effects, such that effect sizes were larger for cohorts of participants who received payment, and for those who received the intervention in a laboratory with an experimenter present. Furthermore, they found that the placebo conditions that used a sham training also showed reductions in bias scores after "training". Cristea and colleagues argued that this reduction in bias in sham conditions is evidence of practice effects of the training tasks, and thus, the tasks may be capturing improvement in task performance rather than a genuine reduction in bias.

Moreover, Cristea and colleagues (2015) found evidence of a time-lag effect, where, over time, effect sizes in studies become smaller. This is a common phenomenon in intervention research where the first studies within a field generally show a large effect to get published in the first place, and this effect becomes moderated as more studies are conducted and unsuccessful replications are published. Cristea and colleagues concluded that the praise for CBM as a potentially effective treatment for clinical disorders had been premature, and that "highly laudatory narrative reviews, comments and editorials, published before the efficiency of the new interventions had been established in well-powered, methodologically appropriate RCTs" (p.14) had contributed to the time lag effect seen in the current state of the literature.

MacLeod and Clarke (2015; see also Clarke et al., 2014) and to a lesser extent, Kuckertz and Amir (2015) have argued that the discrepancies in findings regarding the efficacy of ABM are the result of a failure to change the bias during treatment delivery, rather than a failure of ABM itself. They reported that when training has successfully shifted attention away from the previously capturing stimulus, it has also reliably reduced anxiety vulnerability and symptomatology. However, when no change in symptoms is observed, they argue that the bias was not successfully changed, and thus it was evidence of a fault in the delivery of ABM, not of the inefficacy of ABM itself. MacLeod and Clarke suggested that there may be individual differences in the malleability of attentional bias, which may contribute to the mixed success in training, and argue that research needs to focus on identifying the cognitive mechanisms that are responsible for individual differences in the malleability to attentional change.

Issues Surrounding the use of the Dot Probe Task to Modify Attentional Bias

Whether it be applied to anxiety or depression (or plausibly, any other disorder), the issues surrounding the efficacy of attentional bias modification are the same. We do not understand the mechanism of bias change, nor are we sure that the primary task used to assess this change reliably measures changes in attentional bias. The lack of research on the psychometric properties of the dot probe task was recognized as early as 2005 (Schmukle, 2005), and subsequent research has further called these properties into question (Cisler et al., 2009). The dot probe task is the most popular tool used to assess and modify attentional bias, as well as to measure the success of attentional bias modification. However, without adequate psychometric rigour, it stands that the data derived from this task can only be presented with a level of confidence on par with the properties of the task.

In treatment studies, reliability of a measure is a particularly important factor. The reliability of the measure used to assess outcomes is required to be much higher than for research identifying statistical association or differences between groups (Rodebaugh et al., 2016). This is a critical issue, because the internal consistency (split half and Cronbach's alpha) and test-retest reliability of the dot probe task are inadequate (see Schmukle, 2005). If a measure is both internally inconsistent

and unstable over time, it is unreliable. This raises a very important question: If the dot probe task is not a valid measure of attentional bias, what if anything does it measure and what if anything does attentional bias modification train? Schmukle (2005) reported the existence of a curvilinear relationship between anxiety and bias, whereby low and moderately anxious subjects do not show an attentional bias but highly anxious individuals do (e.g., Broadbent & Broadbent, 1988). However, they also noted the lack of replication of attentional bias for high trait anxious subjects, calling into question even this relationship (e.g., Mogg et al., 1997; Mogg, Millar, & Bradley, 2000). If the dot probe task is unable to reliably detect group differences, the validity of the task cannot be assured.

Furthermore, Price and colleagues (2015) noted that if test–retest reliability for a specific index of bias has not been established, statistical tests for changes in this bias may be invalid. Therefore, as the dot probe task has low test-retest reliability, the ability of the task to effectively capture changes in bias resulting from ABM may also be problematic. Price and colleagues further argued that reliability may vary as a function of sample characteristics, task design, and analysis decisions made by the experimenter. For example, stimuli positioned vertically produce stronger attentional bias than stimuli positioned horizontally, and stimulus-type (pictures of emotional faces versus written words versus abstract/other pictorial stimuli) moderated the strength of the bias (Beard et al., 2012). Differences in task protocol may be responsible for the disparate findings; however, until there is a standardised protocol for administering the dot probe task as an assessment or as a training tool, evaluation of treatment efficacy remains difficult.

Rodebaugh et al. (2016), in discussing attentional bias toward threat, noted that most studies that use the dot probe task do not report psychometric properties of the attentional bias measure, and those that do typically report poor reliability. Two primary methodological worries arise from poor reliability of the attentional bias measure. These concern the sensitivity and specificity of participant selection based on individual differences, and the validity of mediation analyses linking attentional bias to clinical disorders via a third variable (Rodebaugh et al., 2016).

Sensitivity and specificity of a measure are important when deciding who is (and is not) suitable to be targeted by a treatment, and even more so when assessing the efficacy of that treatment. Rodebaugh et al. (2016) report that internal consistency of at least .90 is needed to be confident that decisions about individuals — for example, decisions about whether a certain individual is a good candidate for ABM, or whether treatment has been successful in shifting bias— are accurate. Thus, confidence in the apparent success of modification training is limited by the reliability of the measure. To date, measures of internal consistency of the attentional bias score from the dot-probe task fall well below the recommended level of at least .90, hovering around .45 (see Rodebaugh et al., 2016, for a review).

Additionally, without adequate reliability of the bias measure, confidence in concluding that a third variable (i.e., the modification training) was the mechanism for change is compromised. Specifically, if the reliability of the attentional bias measure cannot be established, alternative explanations for a shift in bias scores, such as random error or regression to the mean, could be responsible for changes in scores across time (Rodebaugh et al., 2016). Hence, adequate reliability is important to draw firm conclusions as to the efficacy of ABM.

MacLeod and Clarke (2015) argued that the dot probe task can reliably detect group differences (e.g., clinical versus non-clinical; Bar-Haim et al., 2007) as well as the success of ABM training conditions (e.g., active versus placebo training; Hakamata et al., 2010); however, they acknowledge that the dot probe task is not a satisfactory measure of inter-individual differences. Price and colleagues (2015) also acknowledged the instances where the dot probe task could differentiate between groups but noted the published exceptions where no group differences were found (e.g., Mohlman, Price, & Vietri, 2013; Price et al., 2013; Waters, Lipp, & Spence, 2004). They further highlighted the risk of file drawer effect, which means that there are likely numerous studies that did not find differences and thus were not published, emphasizing the necessity for pre-registration of studies.

It may be that the dot probe task is not sensitive enough to detect attentional bias. For example, Mogg and Bradley (2005) found an attentional bias using an eye-tracking task, but not using the dot probe task. And even when it does detect the attentional bias, the dot probe task can only reveal that something has happened, not what has happened or when (Cisler et al., 2009). As a result, other ways of analysing the data from the dot probe task have been recommended to improve the task's test-retest reliability. In studies where stimuli are vertically aligned, Kuckertz and Amir (2015) recommended using only data from the bottom dot-location to reduce noise, as the bottom location appears to capture bias more strongly than the top location. Alternatively, Price and colleagues (2015) recommended using only horizontal stimuli allocation to reduce the noise from the vertical alignment positional bias. They also recommended combining multiple assessment points into one measure, as results from numerous dot probe assessments are more stable than results from a single assessment. They further suggested using a bias variability index rather than a mean bias score, as the variability index has better test-retest reliability. As a final point, they discussed the option of Winsorizing data (rescaling outliers) or adopting more comprehensive analysis techniques to enable reliable inferences about individuals to be made (Price et al., 2015). In response to this last suggestion, we would not advocate for Winsorizing data because by cleaning data to improve reliability we risk discarding valuable information. That is, cognitive effects of interest might well manifest in the tail end of the RT distribution (Heathcote, Popiel, & Mewhort, 1991). As an alternative, analysis by fitting an evidence accumulation model to data from the dot probe task offers the possibility of discerning specific decisional processes of interest.

Way Forward: Evidence Accumulation Models

Evidence accumulation models such as the diffusion decision model (DDM; Ratcliff & McKoon, 2008) and the linear ballistic accumulator model (LBA; Brown & Heathcote, 2008) isolate and quantify components of cognition that underlie speeded decision-making, decomposing behavioural data - accuracy, mean response times, and response time distributions - into meaningful components of cognitive processing. The components, or parameters, that evidence accumulation

models return represent well-defined and theoretically meaningful psychological constructs. The models posit that speeded decision-making can be conceptualised as a process in which evidence accrues continuously from a starting point to a decision threshold, at which point a response is triggered. Some models, e.g. the DDM, assume that evidence accrues along a single bipolar scale, with thresholds corresponding to two different response options lying at opposite ends of the scale response. Other models, e.g. the LBA model, assume a separate accumulator corresponding to each potential response. Focussing on the diffusion decision model (see *Figure 1)*, the parameters of central interest are boundary separation, starting point, drift rate, and non-decision time (see Voss, Rothermund & Voss, 2004; Voss & Voss, 2007; or Ratcliff & McKoon, 2008 for a detailed explanation of how behavioural data is transformed into these components).

There were two ways data from the dot probe task can be mapped to the diffusion model. In most circumstances, the model is mapped to the response keys. The response keys for the dot-probe task are generally left/right, or up/down, depending on the configuration of the task. Therefore, the upper and lower response thresholds of the diffusion model represent each of the available response choices (i.e., left/right, up/down). Mapping the data in this manner means that the parameters are based on evidence for the probe appearing on the left or right, or for the arrow pointing up and down. To examine the impact of the stimuli on decision making, parameters would need to be compared based on trial type (congruent/incongruent) and emotion (happy/sad/angry). Differences based on trial type and emotion would then offer insight into how the emotive stimuli impacted the implicit decision-making parameters separated by the diffusion model.

An alternative way to map data from the dot probe task to the diffusion model is to map the response thresholds by congruence (Voss, Voss, & Lerche, 2015). For example, congruent responses are represented by the upper threshold, and incongruent responses are represented by the lower threshold. Mapping the data this way enables direct feedback of the impact of trial type (congruent/incongruent), such that the parameters directly reflect bias toward or away from congruent stimulus/dot pairings. This is based on the premise that if attentional bias is consistent,

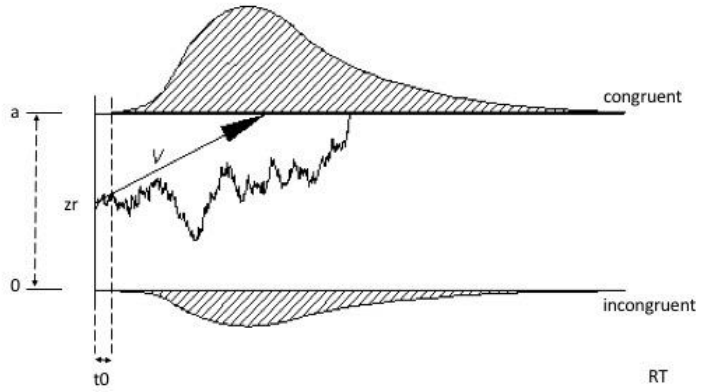


Figure 1. Diffusion decision model response parameters, boundary separation (*a*), relative starting point (z_r), drift rate (*v*), non-decision time (*t0*) for a decision where the response alternatives are congruent and incongruent, as per the dot-probe task. (Diagram adapted from Voss & Voss, 2007; Reproduced with permission.)

when negative information is displayed, attention should be captured by the negative information (Rodebaugh et al., 2016).

Mapping the parameters to the task in this format means that boundary separation (*a*) represents how much evidence is required of the existence of the probe before responding, and start point (z_r) represents an expectation for the probe to appear more on congruent or incongruent trials. Drift rate (*v*) represents the rate at which evidence accumulates for congruent or incongruent responses. If evidence accumulates more quickly for congruent responses, this will result in a positive drift rate, whereas if evidence accumulates more quickly for incongruent responses, this results in a negative drift rate. As non-decisional times (*t0*) represent the speed of processes outside the decisional components of responding, they are theoretically not affected by how the parameters are mapped.

More specifically, *boundary separation (a)* reflects the decision-maker's response caution. This is reflected in the separation between a lower limit decision boundary (0) and an upper decision boundary (*a*), each of which corresponds to one of the two available response choices. Because the value of the lower boundary is fixed at 0, a larger value of *a* indicates a larger boundary separation. A larger boundary separation, holding all other parameters equal, requires more evidence to trigger one of the available response options, and thus results in slower, but more accurate, decision-making, i.e., increased caution.

Relative starting point (z_r) is the point between the decision boundaries from which evidence accumulation starts, measured as a proportion of boundary separation. This parameter reflects the decisional bias an individual has for one response option over another. The closer the starting point is to either boundary, the shorter the processing time will be for the corresponding response, and thus a decisional bias for that response.

Drift rate (v) is the mean rate of evidence accumulation for the different response possibilities. Faster accumulation of evidence, with all other parameters being equal, implies a faster accrual of evidence to the decision threshold, and thus a shorter response time. The rate of evidence accumulation is largely dependent on the physical quality, familiarity, or semantic clarity of the stimulus, and on the skill (Dutilth, Vandekerckhove, Tuerlinckx, & Wagenmakers, 2009), physiological state (Ratcliff & Dongen, 2009), and attentional focus (Ratcliff & Strayer, 2014) of the participant. Therefore, regarding the dot probe task, if the dot is behind a stimulus that has captured attention, an individual may respond more quickly as the evidence accumulation for that location would be faster.

Finally, there is *non-decision time (t0),* which consists of pre-decisional encoding processes (e.g., transmission of neural signals from the retina to the visual cortex) and post-decisional response processes (e.g., execution of a motor response). The non-decision component can identify between-subjects' differences in processes outside of the evidence accumulation. For example, individuals with depression are believed to have motor slowing, so their non-decision time may be

longer than that of individuals without depression. Together, these four components (boundary separation, starting point, drift rate, and non-decision time) determine the RTs and error rates.

Work from Ratcliff, Thapar, and McKoon (2006) provides an elegant demonstration of the way that evidence accumulation models can inform our theoretical understanding of cognitive processes. Older adults are generally slower on speeded decision tasks than young adults, and this difference has often been interpreted as evidence of generalised age-related slowing in cognitive processes (e.g. Myerson, Ferraro, Hale, & Lima, 1992). However, contrary to this suggestion, evidence accumulation models have suggested that older adults' long RTs are the result of high response threshold and slowing of non-decisional processes, rather than to a reduction of evidence accumulation rates (Ratcliff et al., 2006). That is, older adults may not be slower in their cognitive processing than young adults, they may just be more cautious, and slower in their sensory encoding or motor execution. Thus, by separating out the decisional components involved in fast decision-making, Ratcliff and colleagues challenged and advanced the understanding of cognitive aging.

Recently, White, Ratcliff, Vasey, and McKoon, (2010a) used the DDM (Ratcliff & McKoon, 2008) to demonstrate the benefits of using an accumulator model to analyse clinical data over and above analysis of RTs alone. Specifically, they examined behavioural data from a lexical decision task for three separate samples consisting of high and low trait anxious individuals. The data showed only weak, nonsignificant trends hinting at a threat bias for anxious participants when analysed using RTs alone. When these data were analysed with the DDM, White and colleagues found a clear threat bias for high-anxious participants that was not identified by traditional analysis. Thus, the evidence accumulation model identified effects that analysis of mean RTs overlooked.

Traditional dot probe analysis uses only correct RTs, whereas evidence accumulation models use RT distributions for both correct and error responses, localising differences in RTs and/or accuracy to the component(s) of processing that are responsible. This statistical separation of the parameters reduces variance and provides better statistical sensitivity than traditional analyses of mean RTs, which conflate variance from multiple underling cognitive processes (White et al.,

2010a; White, Ratcliff, Vasey, & McKoon, 2010b). Further, as demonstrated by White and colleagues (2010a) using the data from a lexical decision task, evidence accumulation models can identify processing differences even when there are no apparent differences in the behavioural data. This is particularly relevant to the mixed results obtained from the dot probe task. Within an evidence accumulation model, attentional bias may mean any of three things: (1) individuals with a negative attentional bias adopt a lower threshold for interpreting something as a threat or as dysphoric, or for responding to threats and dysphoric stimuli, than do people without the bias; (2) individuals with negative attentional bias accumulate evidence more quickly from negative stimuli—threats in the case of individuals with anxiety, dysphoric items in the case of individuals with depression—than those who do not have a bias, i.e., they find negative stimuli more engaging and/or stronger in intensity; or (3) individuals with negative attentional bias show shorter non-decision times—most likely due to faster motor execution—for processing negative items than do non-biased individuals.

Notably, Mathews and McLeod posited in 2005 that attention to threat seems to depend on some intensity or urgency threshold. Below this threshold danger signals can be ignored but above it they are actively attended. This hypothesis, which in effect interprets negative attentional bias as a higher drift rate, is one that evidence accumulation models can test directly.

Clinical Implications

The clinical benefit of analysing data from the dot probe task using evidence accumulation models is two-fold. First, it would enable us to identify the decisional component or components that are responsible for an attentional bias for negative information. In doing so, we would be able to determine whether there is a specific, implicit cognitive profile that represents attentional bias trans-diagnostically, or rather whether there are profiles that are unique to different disorders. Second, the use of evidence accumulation models would enable us to identify which decisional component or components are targeted by the modified dot probe, and which decisional components underlie successful attentional change. This would enable us to identify those

individuals for whom ABM would be most successful. In addition, if there were any decisional components, as evident from an attentional bias profile, that the dot probe training task does not target, we could move to develop tasks that do target those components.

Furthermore, the use of ABM as a preventative tool could become an option. If we were able to identify the implicit cognitive profile of individuals with an attentional bias for negative information, ABM may provide a useful tool for modifying the bias prior to the development of symptoms. By identifying the specific cognitive profile/s involved in attentional bias and its modification, we would be able to tailor treatment more effectively to the individual.

Implicit mechanisms are hard to target by current explicit treatments (Mobini et al., 2012). Therefore, a treatment that targets the implicit causal and maintenance mechanisms of anxiety and depression has the potential to be a stand-alone treatment, or a beneficial adjunct therapeutic approach. CBT relies on the individual's capacity for strategic and voluntary processing. An implicit processing style characterised by a strong attentional bias toward negative affective information may interfere with an individual's ability to maintain a more neutral or positive focus that is needed to effectively engage in therapy. Successful ABM therefore has the potential to make traditional CBT sessions more constructive.

Conclusion

Anxiety and depression have the highest prevalence rates of all mental health disorders, and have high rate of relapse. An implicit attentional preference for disorder-specific negative emotional stimuli has been hypothesised in the aetiology and maintenance of both disorders. As a result, a new therapeutic approach that targets this attentional bias, known as attentional bias modification (ABM), has been devised. Some research findings indicate that ABM has therapeutic effects on the symptoms of both anxiety and depression. However, questions remain as to the ability of the primary measure of attentional bias, the dot probe task, to reliably detect an attentional bias, and any changes in this bias. In this paper, we have argued that before discounting the dot probe task as unreliable, evidence accumulation models be used to analyse the data from this task to yield a more

comprehensive interpretation than traditional methods allow. Evidence accumulation models have the capacity to identify the implicit decisional components that underpin an attentional bias to negative information and can identify differences in these decisional components even when there are no apparent differences in the behavioural data. This information can offer novel insights into the underlying implicit cognitive mechanisms of attentional bias and its successful modification. An advanced understanding of attentional bias in anxiety and depression, as well as its mechanisms for change, would enable clinicians to tailor ABM to individual client differences. In so doing, ABM has potential as an effective preventative measure, stand-alone treatment, or a tool to complement current CBT practices.

References

Amir, N., Beard, C., Taylor, C. T., Klumpp, H., Elias, J., Burns, M., & Chen, X. (2009). Attention Training in Individuals with Generalized Social Phobia: A Randomized Controlled Trial. *Journal of Consulting and Clinical Psychology*, *77*(5), 961–973. http://doi.org/10.1037/a0016685

Baert, S., Koster, E. H. W., & De Raedt, R. (2011). Modification of Information-Processing Biases in Emotional Disorders: Clinically Relevant Developments in Experimental Psychopathology. *International Journal of Cognitive Therapy*, *4*(2), 208–222. http://doi.org/10.1521/ijct.2011.4.2.208

Bantin, T., Stevens, S., Gerlach, A. L., & Hermann, C. (2016). What does the facial dot-probe task tell us about attentional processes in social anxiety? A systematic review. *J Behav Ther Exp Psychiatry*, *50*, 40–51. http://doi.org/10.1016/j.jbtep.2015.04.009

Bar-Haim, Y. (2010). Research Review: attention bias modification (ABM): a novel treatment for anxiety disorders. *Journal of Child Psychology and Psychiatry*, *51*(8), 859–870. http://doi.org/10.1111/j.1469-7610.2010.02251.x

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. *Psychol Bull*, *133*(1), 1–24. http://doi.org/10.1037/0033-2909.133.1.1

Beard, C. (2011). Cognitive bias modification for anxiety: current evidence and future directions. *Expert Review of Neurotherapeutics*, *11*(2), 299–311.

Beard, C., Sawyer, A. T., & Hofmann, S. G. (2012). Efficacy of Attention Bias Modification Using Threat and Appetitive Stimuli: A Meta-Analytic Review. *Behavior Therapy*, *43*(4), 724–740. http://doi.org/http://dx.doi.org/10.1016/j.beth.2012.01.002

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. International Universities Press.

Beck, A. T., Emery, G., & Greenberg, R. L. (1985). Anxiety disorders and phobias: A cognitive approach. *Basic, New York*, b58.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). Cognitive therapy of depression. 1979. *New York: Guilford Press Google Scholar*.

Bradley, B. P., Mogg, K., & Lee, S. C. (1997). Attentional biases for negative information in induced and naturally occurring dysphoria. *Behaviour Research and Therapy*, *35*(10), 911–927.

Broadbent, D., & Broadbent, M. (1988). Anxiety and attentional bias: State and trait. *Cognition & Emotion*, *2*(3), 165–183.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*(3), 153–178. http://doi.org/10.1016/j.cogpsych.2007.12.002

Browning, M., Holmes, E., & Harmer, C. (2010). The modification of attentional bias to emotional information: A review of the techniques, mechanisms, and relevance to emotional disorders. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(1), 8–20. http://doi.org/10.3758/CABN.10.1.8

Carver, C. S., Johnson, S. L., & Joormann, J. (2008). Serotonergic function, two-mode models of self-regulation, and vulnerability to depression: what depression has in common with impulsive aggression. *Psychological Bulletin*, *134*(6), 912.

Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016). Scaling-up treatment of depression and anxiety: a global return on investment analysis. *The Lancet Psychiatry*, *3*(5), 415–424. http://doi.org/10.1016/S2215-0366(16)30024-4

Cisler, J. M., Bacon, A. K., & Williams, N. L. (2009). Phenomenological Characteristics of Attentional Biases Towards Threat: A Critical Review. *Cognit Ther Res*, *33*(2), 221–234. http://doi.org/10.1007/s10608-007-9161-y

Cisler, J. M., & Koster, E. H. W. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clinical Psychology Review*, *30*(2), 203–216. http://doi.org/10.1016/j.cpr.2009.11.003

Clarke, P. J., Notebaert, L., & MacLeod, C. (2014). Absence of evidence or evidence of absence: reflecting on therapeutic implementations of attentional bias modification. *BMC Psychiatry*, *14*(1), 1.

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

Dutilh, G., Vandekerckhove, J., Tuerlinckx, F., & Wagenmakers, E. J. (2009). A diffusion model decomposition of the practice effect. *Psychonomic Bulletin & Review*, *16*(6), 1026-36. doi:10.3758/16.6.1026

De Raedt, R., & Koster, E. H. (2010). Understanding vulnerability for depression from a cognitive neuroscience perspective: A reappraisal of attentional factors and a new conceptual framework. *Cognitive, Affective, & Behavioral Neuroscience, 10*(1), 50–70.

Donaldson, C., Lam, D., & Mathews, A. (2007). Rumination and attention in major depression. *Behaviour Research and Therapy*, *45*(11), 2664–2678. http://doi.org/10.1016/j.brat.2007.07.002

Emmelkamp, P. M. (2012). Attention bias modification: the Emperor's new suit? *BMC Medicine*, *10*(1), 63.

Everaert, J., Koster, E. H., & Derakshan, N. (2012). The combined cognitive bias hypothesis in depression. *Clinical Psychology Review*, *32*(5), 413–424.

Fox, E., Russo, R., & Georgiou, G. A. (2005). Anxiety Modulates the Degree of Attentive Resources Required to Process Emotional Faces. *Cognitive, Affective & Behavioral Neuroscience*, *5*(4), 396–404.

Garner, M., Mogg, K., & Bradley, B. P. (2006). Orienting and maintenance of gaze to facial expressions in social anxiety. *Journal of Abnormal Psychology*, *115*(4), 760.

Goodman, S. (Ed.). (2009). Cogntive Bias Modification in Emotional Disorders [Special Section]. *Journal of Abnormal Psychology*, *118*(1).

Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: current status and future directions. *Annual Review of Clinical Psychology*, *6*, 285.

Gotlib, I. H., Krasnoperova, E., Yue, D. N., & Joormann, J. (2004). Attentional biases for negative interpersonal stimuli in clinical depression. *Journal of Abnormal Psychology*, *113*(1), 127.

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*(11), 982–990. http://doi.org/http://dx.doi.org/10.1016/j.biopsych.2010.07.021

Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological Bulletin*, *137*(6), 940.

Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin*,*109*(2), 340.

Hertel, P. T., & Mathews, A. (2011). Cognitive Bias Modification: Past Perspectives, Current Findings, and Future Applications. *Perspectives on Psychological Science*, *6*(6), 521–536. http://doi.org/10.1177/1745691611421205

Koster, E. H., Crombez, G., Verschuere, B., & De Houwer, J. (2004). Selective attention to threat in the dot probe paradigm: Differentiating vigilance and difficulty to disengage. *Behaviour Research and Therapy*, *42*(10), 1183–1192.

Koster, E. H., Crombez, G., Verschuere, B., Van Damme, S., & Wiersema, J. R. (2006). Components of attentional bias to threat in high trait anxiety: Facilitated engagement, impaired disengagement, and attentional avoidance. *Behaviour Research and Therapy*, *44*(12), 1757–1771.

Koster, E. H., Verschuere, B., Crombez, G., & Van Damme, S. (2005). Time-course of attention for threatening pictures in high and low trait anxiety. *Behaviour Research and Therapy*, *43*(8), 1087–1098.

Kuckertz, J. M., & Amir, N. (2015). Attention bias modification for anxiety and phobias: Current status and future directions. *Current Psychiatry Reports*, *17*(2), 1–8.

Leppänen, J. M. (2006). Emotional information processing in mood disorders: a review of behavioral and neuroimaging findings. *Current Opinion in Psychiatry*, *19*(1), 34–39.

Leyman, L., De Raedt, R., Schacht, R., & Koster, E. H. W. (2007). Attentional biases for angry faces in unipolar depression. *Psychological Medicine*, *37*(3), 393–402. http://doi.org/10.1017/S003329170600910X

MacLeod, C. (2012). Cognitive bias modification procedures in the management of mental disorders. *Current Opinion in Psychiatry*, *25*(2), 114–120 10.1097/YCO.0b013e32834fda4a.

MacLeod, C., Campbell, L. W., Rutherford, E. M., & Wilson, E. (2004). The causal status of anxiety-linked attentional and interpretive bias. Presented at the Cambridge University Press. Retrieved from http://research-repository.uwa.edu.au/en/publications/the-causal-status-of-anxietylinked-attentional-and-interpretive-bias(5ab5a77e-4856-4e76-b25d-5ab68be66dd6)/export.html

MacLeod, C., & Clarke, P. J. F. (2015). The Attentional Bias Modification Approach to Anxiety Intervention. *Clinical Psychological Science*, *3*(1), 58–78. http://doi.org/10.1177/2167702614560749

MacLeod, C., & Mathews, A. (2012). Cognitive bias modification approaches to anxiety. *Annual Review of Clinical Psychology*, *8*, 189–217.

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, *95*(1), 15.

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, 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*(1), 107.

Mathews, A., & Mackintosh, B. (1998). A Cognitive Model of Selective Processing in Anxiety. *Cognitive Therapy and Research*, *22*(6), 539–560. http://doi.org/10.1023/A:1018738019346

Mathews, A., & MacLeod, C. (1985). Selective processing of threat cues in anxiety states. *Behaviour Research and Therapy*, *23*(5), 563–569.

Mathews, A., & MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annu. Rev. Clin. Psychol.*, *1*, 167–195.

Mobini, S., Reynolds, S., & Mackintosh, B. (2012). Clinical Implications of Cognitive Bias Modification for Interpretative Biases in Social Anxiety: An Integrative Literature Review. *Cognitive Therapy and Research*, *37*(1), 173–182. http://doi.org/10.1007/s10608-012-9445-8

Mogg, K., Bradley, B., Miles, F., & Dixon, R. (2004). Brief report time course of attentional bias for threat scenes: testing the vigilance-avoidance hypothesis. *Cognition and Emotion*, *18*(5), 689–700.

Mogg, K., & Bradley, B. P. (2005). Attentional bias in generalized anxiety disorder versus depressive disorder. *Cognitive Therapy and Research*, *29*(1), 29–45.

Mogg, K., Bradley, B. P., De Bono, J., & Painter, M. (1997). Time course of attentional bias for threat information in non-clinical anxiety. *Behaviour Research and Therapy*, *35*(4), 297–303.

Mogg, K., Bradley, B. P., & Williams, R. (1995). Attentional bias in anxiety and depression: the role of awareness. *The British Journal of Clinical Psychology / the British Psychological Society*, *34 (Pt 1)*, 17–36.

Mogg, K., Mathews, A., & Weinman, J. (1987). Memory bias in clinical anxiety. *Journal of Abnormal Psychology*, *96*(2), 94–98. http://doi.org/10.1037/0021-843X.96.2.94

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*(4), 695.

Mogoaşe, C., David, D., & Koster, E. H. (2014). Clinical efficacy of attentional bias modification procedures: An updated meta-analysis. *Journal of Clinical Psychology*, *70*(12), 1133–1157.

Mohlman, J., Price, R. B., & Vietri, J. (2013). Attentional bias in older adults: Effects of generalized anxiety disorder and cognitive behavior therapy. *Journal of Anxiety Disorders*, *27*(6), 585–591. http://doi.org/10.1016/j.janxdis.2013.06.005

Myerson, J., Ferraro, F. R., Hale, S., & Lima, S. D. (1992). General slowing in semantic priming and word recognition. *Psychology and Aging*, *7*(2), 257.

NICE. (2011). Common mental health problems: identification and pathways to care. Clinical guideline 123. Retrieved September 18, 2016, from https://www.nice.org.uk/Guidance/CG123

Peckham, A. D., McHugh, R. K., & Otto, M. W. (2010). A meta-analysis of the magnitude of biased attention in depression. *Depression and Anxiety*, *27*(12), 1135–1142.

Pergamin-Hight, L., Naim, R., Bakermans-Kranenburg, M. J., van, Ij. M. H., & Bar-Haim, Y. (2015). Content specificity of attention bias to threat in anxiety disorders: a meta-analysis. *Clin Psychol Rev*, *35*, 10–8. http://doi.org/10.1016/j.cpr.2014.10.005

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25.

Price, R. B., Kuckertz, J. M., Siegle, G. J., Ladouceur, C. D., Silk, J. S., Ryan, N. D., … Amir, N. (2015). Empirical recommendations for improving the stability of the dot-probe task in clinical research. *Psychol Assess*, *27*(2), 365–76. http://doi.org/10.1037/pas0000036

Price, R. B., Siegle, G. J., Silk, J. S., Ladouceur, C., McFarland, A., Dahl, R. E., & Ryan, N. D. (2013). Sustained neural alterations in anxious youth performing an attentional bias task: a pupilometry study. *Depression and Anxiety*, *30*(1), 22–30. http://doi.org/10.1002/da.21966

Ratcliff, R., & Dongen, H. P. A. (2009). Sleep deprivation affects multiple distinct cognitive processes. *Psychonomic Bulletin and Review*, *16*(4), 742–751. http://doi.org/10.3758/PBR.16.4.742

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.

Ratcliff, R., Thapar, A., & McKoon, G. (2006). Aging and individual differences in rapid two-choice decisions. *Psychonomic Bulletin & Review*, *13*(4), 626–635.

Ratcliff, R., & Strayer, D. (2014). Modeling simple driving tasks with a one-boundary diffusion model. *Psychonomic Bulletin & Review*, *21*(3), 577-89. doi:10.3758/s13423-013-0541-x

Rodebaugh, T. L., Scullin, R. B., Langer, J. K., Dixon, D. J., Huppert, J. D., Bernstein, A., … Lenze, E. J. (2016). Unreliability as a threat to understanding psychopathology: The cautionary tale of attentional bias. *Journal of Abnormal Psychology*, *125*(6), 840–851. https://doi.org/10.1037/abn0000184

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Vol. 6). Sage.

Salemink, E., van den Hout, M. A., & Kindt, M. (2007). Selective attention and threat: Quick orienting versus slow disengagement and two versions of the dot probe task. *Behaviour Research and Therapy*, *45*, 607–615.

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*(1), 5.

Schmukle, S. C. (2005). Unreliability of the dot probe task. *European Journal of Personality*, *19*(7), 595–605.

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*(1), 65.

Shechner, T., Britton, J. C., Pérez-Edgar, K., Bar-Haim, Y., Ernst, M., Fox, N. A., … Pine, D. S. (2012). Attention biases, anxiety, and development: toward or away from threats or rewards? *Depression and Anxiety*, *29*(4), 282–294.

Staugaard, S. R. (2009). Reliability of two versions of the dot-probe task using photographic faces. *Psychology Science Quarterly*, *51*(3), 339–350.

Van Bockstaele, B., Verschuere, B., Tibboel, H., De Houwer, J., Crombez, G., & Koster, E. H. (2014). A review of current evidence for the causal impact of attentional bias on fear and anxiety. *Psychological Bulletin, 140*(3), 682.

Voss, A., & Voss, J. (2007). Fast-dm: a free program for efficient diffusion model analysis. *Behavior Research Methods, 39*(4), 767–775.

Voss, A., Voss, J., & Lerche, V. (2015). Assessing cognitive processes with diffusion model analyses: a tutorial based on fast-dm-30. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00336

Waters, A. M., Lipp, O. V., & Spence, S. H. (2004). Attentional bias toward fear-related stimuli: an investigation with nonselected children and adults and children with anxiety disorders. *Journal of Experimental Child Psychology, 89*(4), 320–337. http://doi.org/10.1016/j.jecp.2004.06.003

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*(1), 55.

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010a). Anxiety enhances threat processing without competition among multiple inputs: a diffusion model analysis. *Emotion (Washington, D.C.), 10*(5), 662–677. http://doi.org/10.1037/a0019474

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010b). Using diffusion models to understand clinical disorders. *Journal of Mathematical Psychology, 54*(1), 39–52. http://doi.org/http://dx.doi.org/10.1016/j.jmp.2010.01.004

Chapter 3: Study 1

Reliability of the attentional bias score: A diffusion model approach to analysis

Amy Manuel, Eva Kemps, Jason S. McCarley

School of Psychology, Flinders University, Adelaide, Australia

Author Note

Jason S McCarley, School of Psychology, Flinders University, is now at the College of Liberal Arts, Oregon State University, USA

Correspondence concerning this article should be addressed to Amy Manuel, School of Psychology, Flinders University, GPO Box 2100, Adelaide, SA 5001, Australia. Electronic mail may be sent to Amy.Manuel@flinders.edu.au

Abstract

The dot probe task is the primary method used to investigate the link between attentional bias for mood congruent stimuli and anxiety and depression. However, the reliability of the attentional bias score derived from mean RTs falls well below recommended levels for clinical research. In this paper, we applied a more sophisticated model of analysis, the Diffusion decision model (DDM), to data from the dot probe task. Our aim was to assess the suitability of DDM parameters as alternative measures of attentional bias. Participants (N = 91) performed a conventional dot probe task with faces as stimuli. While neither form of analysis showed an overall attentional bias, both were related to scores on clinical measures of anxiety, depression, and emotion regulation difficulties. Traditional RT analysis showed that an avoidance of emotion was linked to clinical symptoms; however, by breaking down RTs into parameters, the DDM revealed more specifically that speed of information uptake was linked to clinical symptoms, whereas decision caution was a protective factor. The information returned by the DDM supports attentional control theory, and partially supports the content specificity hypothesis. This is the first study to apply the DDM to data from the dot probe task.

Key words: Attentional bias, Diffusion decision model, Anxiety, Depression, Dot Probe

Reliability of the attentional bias score: A diffusion model approach to analysis

Errors in information processing drive individuals to seek out and attend to mood congruent stimuli, resulting in an attentional bias (Beck, 1976), and an increased difficulty regulating strong emotion (Cisler & Olatunji, 2012; Joormann & Quinn, 2014). Attentional bias is theorised to be a causal factor in the aetiology and maintenance of anxiety and depression (Beck, 1976; Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & Van Ijzendoorn, 2007; Mathews & Mackintosh, 1998; Mathews & MacLeod, 2005; Mogg & Bradley, 2005; Pergamin-Hight, Naim, Bakermans-Kranenburg, Van Ijzendoorn, & Bar-Haim, 2015; Williams, Watts, MacLeod, & Mathews, 1997). Specifically, in line with Beck's *content specificity hypothesis* (Beck, 1976), anxiety has been associated with an implicit attentional preference for threatening information (e.g. Koster, Verschuere, Crombez, & Van Damme, 2005; Mathews & MacLeod, 1985; Pergamin-Hight, Naim, Bakermans-Kranenburg, van, & Bar-Haim, 2015; Shechner et al., 2012), whereas depression has been associated with an implicit attentional preference for dysphoric information (e.g. Gotlib, Krasnoperova, Yue, & Joormann, 2004; Mogg & Bradley, 2005; Peckham, McHugh, & Otto, 2010). By contrast, non-depressed and non-anxious individuals typically prioritize positive over negative emotional information (e.g. Švegar, Kardum, & Polič, 2013).

Similarly, the inability to self-regulate strong emotion is linked with the onset and maintenance of anxiety and depression (Gross, 2013). The ability to regulate strong emotion, either implicitly or explicitly, is an important part of social functioning and good mental health (Gross, 2013). Emotion regulation typically aims to decrease negative emotion and increase positive emotion (Gross, 2013). Emotion regulation processes can be explicit, e.g. consciously trying to calm one's nerves before a performance, or implicit, e.g. unconsciously diverting one's gaze from threatening stimuli (Gross, 2013). For individuals who are unable to effectively self-regulate, attentional bias to disorder-congruent stimuli works to increase experiences of negative emotion, which can exacerbate disorder-relevant symptoms.

The dot probe task is the primary task used to measure biased attentional processes in anxiety and depression (Pergamin-Hight et al., 2015). The dot-probe task was developed by MacLeod and Mathews (1986) based on a paradigm by Posner (1980), and was later modified by Koster and colleagues (2004). But, in recent years, the reliability of the attentional bias score produced by the dot probe task has been questioned (Cisler, Bacon, & Williams, 2009; Schmukle, 2005; Staugaard, 2009), casting doubt on the validity of the measure. The current paper applies a novel form of analysis to data from the dot probe task. Specifically, we applied the diffusion model (Ratcliff & McKoon, 2008), a mathematical model of choice response time (RT), to ascertain if this was a more robust method of analysing data from the dot probe task.

Measurement of Attentional Bias

To understand reliability issues that surround attentional bias measurement it is important to first understand how the measure of attentional bias is derived from the dot-probe task. The task asks participants to focus on a fixation cross in the centre of a computer screen. The fixation interval is followed by the brief presentation of two stimuli, typically images or words, side-by-side or vertically aligned. One stimulus is emotional and the other is neutral in valence. After the stimuli disappear, one is replaced by a dot, and the participant is required to execute a speeded response indicating which of the two potential positions the probe appeared in (Koster, Crombez, Verschuere, & De Houwer, 2004; MacLeod, Mathews, & Tata, 1986). The participant is instructed to ignore the stimuli and focus only on the location of the probe. Trials on which the dot replaces the emotional stimulus are called congruent trials, and trials on which the dot replaces the neutral stimulus are called incongruent trials. If the participant's attention has been captured by a specific stimulus, it is assumed to be reflected by faster RTs when the dot appears in that location. An estimate of attentional bias for emotional stimuli is calculated by subtracting the mean RT for congruent trials from the mean RT for incongruent trials. Positive scores result from faster RTs to the probe behind the emotional stimuli, and are taken to indicate an attentional bias toward that emotion.

The task may also incorporate trials that include only neutral stimuli. These trials, termed non-critical trials, provide a baseline RT measure, with no influence of emotion on attentional capture. Trials that consist of a target-relevant item (e.g., an emotional stimulus) and a neutral item are termed critical trials. Mean RTs for probes appearing behind the emotional stimuli in critical trials (i.e., congruent trials) are subtracted from mean RTs for non-critical trials to calculate a measure of *vigilance*, also referred to as *speeded detection*. Speeded detection is considered a measure of how strongly a target stimulus captures an individual's attention relative to their baseline engagement. Positive scores indicate increased vigilance (Koster et al., 2004), which is referred to as a *detection bias*. Conversely, mean RTs for probes appearing behind the neutral stimulus in critical trials (i.e., incongruent trials) are subtracted from baseline RTs to calculate a measure of *slowed disengagement*. Slowed disengagement is considered a measure of how long it takes an individual to take attention away from the emotional stimulus and redirect it to the probe that appears behind the neutral stimulus. Negative scores indicate increased difficulty disengaging (Koster et al., 2004), or *disengagement bias*. To date, much of the reliability analysis of the dot-probe task has been done with the overall attentional bias score and as such, that will remain our focus in this paper.

Issues of Reliability of the Attentional Bias Score

The reliability of the attentional bias score derived from the dot probe task has become a focus of concern in recent years due to the task being used in treatment studies. Efforts to mitigate the onset, and alleviate symptomology, of anxiety and depression have led researchers to try to modify attentional biases. This approach is known as *cognitive bias modification (CBM)*, and primarily uses the dot-probe task both to shift bias and to measure the success of any shift. However, in treatment studies, reliability of the measure used to assess outcomes is required to be much higher than for research identifying statistical association or differences between groups (Rodebaugh et al., 2016). Consequently, concerns have arisen that the current literature base may be

a biased representation of the efficacy of ABM, such that the effect seems stronger and more robust than it is (Beard, Sawyer, & Hofmann, 2012; Hakamata et al., 2010; Peckham et al., 2010).

Rodebaugh et al. (2016), in discussing attentional bias toward threat, noted that most studies that use the dot probe task do not report psychometric properties of the attentional bias measure, and those that do typically report poor reliability. Two primary methodological worries arise from poor reliability of the attentional bias measure. These concern the sensitivity and specificity of participant selection based on individual differences, and the validity of mediation analyses linking attentional bias to clinical disorders via a third variable (Rodebaugh et al., 2016).

Sensitivity and specificity of a measure are important when deciding who is (and is not) suitable to be targeted by a treatment, and even more so when assessing the efficacy of that treatment. Rodebaugh et al. (2016) report that internal consistency of at least .90 is needed to be confident that decisions about individuals — for example, decisions about whether a certain individual is a good candidate for ABM, or whether treatment has been successful in shifting bias— are accurate. Thus, confidence in the apparent success of modification training is limited by the reliability of the measure. To date, measures of internal consistency of the attentional bias score from the dot-probe task fall well below the recommended level of at least .90, hovering around .45 (see Rodebaugh et al., 2016, for a review).

Additionally, without adequate reliability of the bias measure, confidence in concluding that a third variable (i.e., the modification training) was the mechanism for change is compromised. Specifically, if the reliability of the attentional bias measure cannot be established, alternative explanations for a shift in bias scores, such as random error or regression to the mean, could be responsible for changes in scores across time (Rodebaugh et al., 2016). Hence, adequate reliability is important to draw firm conclusions as to the efficacy of ABM.

Several authors have attempted to improve the reliability of the dot-probe task, from adding more trials (Price et al., 2015), to exploring different stimuli such as words and images (Schmukle, 2005; Staugaard, 2009), to making modifications to the task such as changing stimulus duration

(Schmukle, 2005), to creating new bias scores (Price et al., 2015; Zvielli, Bernstein, & Koster, 2015). The results of these efforts have been mixed, and none have achieved a level of reliability required for a clinical assessment task. Alternative suggestions to address the reliability issues of the attentional bias score derived from the dot-probe task have included (a) finding a new task, (b) being open to the possibility that the theory that attentional bias is a stable construct is flawed, and (c) applying a more sophisticated analysis technique that may have better sensitivity to reliably detect change. It is this last suggestion that we address in this paper.

Traditional analysis techniques for the dot-probe task are based on mean RTs for correct responses, from which inferences are made about how the pattern of results map onto constructs. In addition to issues of reliability surrounding the current bias score, traditional analysis is unable to identify with any certainty the underlying reasons behind the patterns of RTs that are observed. For example, a slow response could indicate that the task is difficult but could also reflect a more cautious decision style. White, Ratcliff, Vasey, and McKoon (2010) postulated that the current analysis of correct mean RTs derived from the dot probe task may not be sensitive enough to reliably detect effects. They suggested that a more comprehensive form of analysis, by way of an *evidence accumulation model*, may be more sensitive due to the model's ability to statistically separate parameters of the decision process.

Evidence accumulation models use a small number of parameters to explain and reproduce the patterns of RTs and error rates seen in speeded-choice decision making. The model parameters represent well-defined and theoretically meaningful psychological constructs. By isolating and quantifying the components of cognition that underlie speeded decision making, evidence accumulation models can provide information about cognitive processes that are hidden from traditional analysis (Voss, Rothermund, & Voss, 2004; Voss & Voss, 2007). One of the most broadly applied of the evidence accumulation models is the *diffusion decision model* (Ratcliff, 1978; Ratcliff & McKoon, 2008; Ratcliff & Smith, 2015). The interpretational validity of the model has been confirmed (Voss, Rothermund & Voss, 2004), as has the ability of the model to robustly

detect moderate changes to the response distribution across trials (Ratcliff, 2013). Ratcliff's diffusion model fits behavioural data as well as or better than competing models (Ratcliff & Smith, 2015, 2004; White et al., 2010).

Evidence accumulation models of analysis can be difficult to fit to data, and as such, until recently, have not received much attention in the analysis of clinical disorders. However, with several methods of fitting the models to data now available to researchers (Fast-dm, Voss & Voss, 2007; EZ diffusion, Wagenmakers, van der Maas, & Grasman, 2007; DMAT, Vandekerckhove & Tuerlinckx, 2008), the models are becoming easier to use and apply. This study applied the diffusion decision model using one of these programs, *Fast-dm 30 (*Voss & Voss, 2007), to data from the dot probe task to assess fit, reliability of parameters across the task, and explore whether the data returned offer any further insight into the cognitive processes of attentional bias.

The Diffusion Decision Model

The diffusion decision model posits that an individual deciding between two options accumulates evidence for one or the other option until a decision threshold is reached and a response is executed. The model transforms behavioural data - accuracy, mean response times, and response time distributions – from two-choice decision making tasks into meaningful components of cognitive processing resulting in four parameters: boundary separation, relative starting point, drift rate and non-decision time (see *Figure 1* and Table 1; For a detailed explanation of how behavioural data is transformed into these components see Voss, Rothermund & Voss, 2004; Voss & Voss, 2007; or Ratcliff & McKoon, 2008).

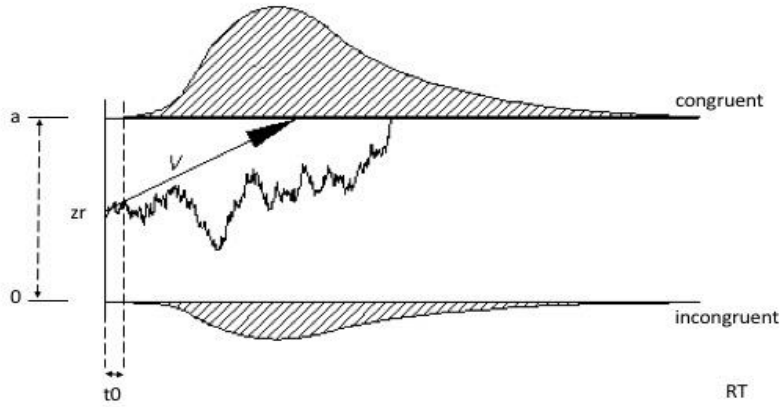


Figure 1. Diffusion decision model response parameters, boundary separation (*a*), relative starting point (z_r), drift rate (*v*), non-decision time (*t0*) for a decision where the response alternatives are congruent and incongruent, as per the dot-probe task. (Diagram adapted from Voss & Voss, 2007; Reproduced with permission.)

Boundary separation *(a)*. The participant's response caution is reflected in the separation between a lower limit decision boundary (0) and an upper decision boundary (a), each of which corresponds to one of the two available choices. Because the value of the lower boundary is fixed at 0, a larger value of *a* indicates a larger boundary separation, meaning that the individual requires more evidence to reach a decision (Pergamin-Hight et al., 2015). The result of a larger boundary separation, holding all else equal, will be responses that are on average slower but more accurate.

Relative starting point (z_r) Relative starting point is the point between the decision boundaries from which evidence accumulation starts, measured as a proportion of boundary separation. This parameter reflects the decisional bias an individual has for one response option over another. The closer the starting point is to either boundary, the shorter the processing time will be for the corresponding response. An unbiased starting point will be returned as 0.5, indicating an equal distance from either decision boundary.

Table 1

Typical parameter values of the drift-diffusion model as returned by fast-dm(Voss, Voss, & Lerche, 2015).

Parameter	Typical range	Description
Boundary separation (*a*)	$0.5 < a < 2$	Larger values equate to a more conservative decision style
Relative Starting Point (z_r)	$0.3 < z_r < 0.7$	Deviations from 0.5 indicate that different amounts of evidence are required to make a decision for the alternative responses.
Drift rate (*v*)	$-5 < v < 5$	The rate of evidence accumulation. Values further from zero indicate stronger evidence. Positive values indicate evidence for the upper threshold response, negative values indicate evidence for the lower threshold response.
Non-decision time (*t0*)	$0.1 < t0 < 0.5$	Average duration of all non-decisional processes (encoding and response execution).

Drift rate (*v*). Drift rate is the mean rate of evidence accumulation toward one choice or the other, that is, the rate of approach to the decision threshold. A higher drift rate means faster accumulation of evidence. Drift rate is generally a function of stimulus or evidence quality. The easier a stimulus is to identify, the higher the drift rate.

Non-decision time (*t0*). The final parameter of relevance is non-decision time, *t0*, the time needed for sensory encoding prior to the evidence accumulation process and for response execution once a decision boundary has been crossed.

Thus, differences in mean RT can be due to faster accumulation of evidence from the stimulus, a lower decision threshold, a biased relative starting point, faster non-decision components, or a combination of the four. Conversely, equal mean RTs may result from different combinations of values for these parameters. By separating the influence of multiple psychological constructs, the analysis of drift diffusion parameters isolates sources of performance variance that are conflated in mean RTs. Therefore, the diffusion decision model has the potential to enhance our understanding of the decision components that are captured by the dot-probe task, and to increase the sensitivity and reliability of attentional bias measures.

Mapping Parameters onto Current Theory

We mapped data from the dot probe task to the diffusion model response thresholds by congruence. For example, congruent responses are represented by the upper threshold, and incongruent responses are represented by the lower threshold. Mapping the data this way enables direct feedback of the impact of trial type (congruent/incongruent), such that the parameters directly reflect bias toward or away from congruent stimulus/dot pairings. This is based on the premise that if attentional bias is consistent, when negative information is displayed, attention should be captured by the negative information, and thus, the equivalent processes in the opposite direction for the other stimulus should shift also (Rodebaugh et al., 2016).

Mapping the parameters to the task in this format means that boundary separation (a) represents how much evidence is required of the existence of the probe before responding, and start point (z_r) represents an expectation for the probe to appear more on congruent or incongruent trials. Drift rate (v) represents the rate at which evidence accumulates for congruent or incongruent responses. If evidence accumulates more quickly for congruent responses, this will result in a positive drift rate, whereas if evidence accumulates more quickly for incongruent responses, this

results in a negative drift rate. As non-decisional times (*t0*) represent the speed of processes outside the decisional components of responding, they are theoretically not affected by how the parameters are mapped.

In modelling data from the dot probe task with Fast-dm 30, we expected that an attentional bias would manifest in the drift rate parameter, *v*. During congruent trials, if attention has been captured by the stimulus, i.e., an attentional bias, the participant's attention will already be deployed to the probe location by the time the probe appears. Therefore, evidence accumulation for the existence of the probe on the corresponding side will be faster than if the probe appears at the opposite, unattended, location. For incongruent trials, we conversely expected that if attention has been captured by the emotional stimulus the rate of evidence accumulation will be slower as it will require the individual to shift attention to the opposite side of the screen. Thus, a bias toward emotional stimuli would be evident from a positive drift rate, whereas an avoidance of emotional stimuli would be evident from a negative drift rate. If, however, there is no difference between the congruent and incongruent drift rates, we surmised that neither the emotional stimuli nor the neutral stimuli captured attention, and thus no bias was evident for that emotion. Given current theories of psychopathology, we expected that drift rate for angry trials (angry facial expression paired with neutral expression) would be linked to scores on anxiety, whereas drift rate for sad trials (sad facial expression paired with neutral expression) would be linked to scores on depression.

Furthermore, it is possible that the non-decision time, *t0*, may vary for different emotions. For example, anxiety has been associated with hypervigilance (Eysenck, Derakshan, Santos, & Calvo, 2007), which may manifest in faster non-decisional processes. Therefore, we might expect to see faster non-decision time to the presentation of threatening stimuli (e.g., angry faces) in individuals with higher anxiety scores. In the same way, depression has been associated with motor slowing (Sobin & Sackeim, 1997); therefore, we might expect to see that non-decision time is longer for sad stimuli in individuals with higher depression scores. Thus, we expected that non-

decision time for angry faces would correlate with anxiety scores, and non-decision time for sad faces would correlate with depression scores.

Finally, as attentional bias has been associated with difficulty regulating strong emotion (Cisler & Olatunji, 2012; Joormann & Quinn, 2014), and emotion regulation difficulty is associated with both anxiety and depression, we expected to see a correlation between the diffusion model parameters and difficulty regulating emotion. Specifically, we hypothesised that drift rates and non-decision time for sad and angry trials would be associated with increased difficulty regulating emotion, whereas drift rate and non-decision time for happy trials (happy facial expression paired with neutral expression) would be associated with better emotion regulation. Further, we expected that smaller boundary separation may be associated with more impulsive decision making, and thus increased emotion regulation difficulties. Due to the equal number of congruent and incongruent trials, we had no theoretical reason to expect starting point to differ between emotions, or to be related to clinical measures. In the same way, due to the presentation of all combinations of critical and non-critical trials in each block of the task, boundary separation cannot vary between emotion as participants had no way of knowing which emotion would be presented next. However, it is plausible that there is a link between overall decision style (i.e., mean boundary separation across all trials) and emotion regulation and anxiety and depression. Both emotion regulation and cautious responding require top down, frontal lobe processing (Etkin, Büchel, & Gross, 2015; Sakagami & Pan, 2007), and this can be impeded in individuals with anxiety and depression (Snyder, 2013; Tovote, Fadok, & Lüthi, 2015). Therefore, we also expected that boundary separation may correlate with emotion regulation, depression and anxiety such that a less conservative boundary setting would be related to higher scores on clinical measures, i.e., poorer frontal lobe activation results in impulsive responding.

Method

This study received ethics approval from the Flinders University Social and Behavioural Research Ethics Committee (Project number 6259: Investigating cognitive mechanisms underpinning emotional cognitive biases in individuals with varying levels of emotional disorders).

Participant Characteristics and Exclusion Criteria

One hundred undergraduate women were recruited through the university's research participation program. Fifty participants received course credit for a first-year psychology topic, and fifty participants received AU$15 remuneration. Seven participants were excluded from the final data set for failing attention checks in the self-report measures, and a further two participants were excluded during the data cleaning process. Thus, data from 91 participants, with a mean age of 22 years (min 17, max 47, median 20), were used in the final analysis.

Design

A 4 (emotion: happy, sad, angry) x 2 (trial type: congruent, incongruent) repeated measures design was used. Dependent measures were RT, and Fast-dm 30 parameters *v* and *t0*. RT and all the Fast-dm 30 parameters (*a, sz, v, t0)* were also used to explore their relationship with scores on the clinical measures of depression, anxiety, stress, and emotion regulation difficulties.

Materials

Facial stimuli were used because emotional images are more salient than words (Bradley, Mogg, & Millar, 2000). Additionally, facial expressions are universally understood (Ekman & Friesen, 1986), whereas words can have different meanings for different people. In line with other research in this area (Arndt & Fujiwara, 2012; Niles, Mesri, Burklund, Lieberman, & Craske, 2013; Paulewicz, Blaut, & Kłosowska, 2012; Nim Tottenham, Hare, & Casey, 2011), we selected facial stimuli from the NimStim database (Tottenham et al., 2009). This database consists of 646 facial expression stimuli across 43 faces of different gender and race. The database has expressions consisting of fearful, happy, sad, angry, surprised, calm, neutral, and disgusted, each with validity ratings (Tottenham et al., 2009). There are two sets of each expression. One set consists of the

expressions with models' mouths open, and the other set consists of expressions with mouths closed. In this study, faces with angry, happy, sad, and neutral expressions were used. They were selected from the set with mouths closed. Three faces were excluded as the models did not present with all four (angry, happy, sad, neutral) expressions. All pictures were presented in a black rectangular frame and cropped to just around the face to remove distinguishing external features, allowing the focus to be on the expression. All images were of equal size (332x458 pixels), resolution (120 pixels/inch), and RGB colour depth (8bpc).

Procedure

Data for the dot-probe task were collected at the same time as data for another study investigating an alternative task, the yes/no task (in preparation). Both tasks were completed on a Dell desktop computer with a standard QWERTY keyboard in one testing session that lasted approximately 45 minutes. The cognitive tasks (dot probe and yes/no task; counterbalanced) were presented first to avoid any mood induction that might result from the question content of the self-report measures. The dot probe task was run using Neurobehavioral Systems Inc. software. The clinical measures were collected in Qualtrics, an online survey program. The clinical measures were presented in the same order for each participant with the Difficulties with Emotion Regulation Scale (DERS) presented first, followed by the Depression, Anxiety and Stress Scale (DASS).

Attentional bias. The dot probe task as modified by Koster and colleagues (2004) to include neutral baseline trials was used. The task consisted of 280 trials in total. For each trial, a central white fixation cross was displayed for 500ms on a black background, followed by a picture pair for another 500ms. Picture pairs comprised either a neutral and emotional (happy, sad, angry) facial expression of the same individual, or two of the same neutral faces. Pictures were presented to the left and right of the fixation cross, centred horizontally, and 6.87cm from both the centre and the perimeter of the screen. A white dot, 3mm in diameter, was presented immediately afterwards in the location previously occupied by either the left or right image, centred horizontally. Presentation of the dot was terminated by the participant pressing a response key to report on which side (left or

right) the dot appeared. Response keys were the Z key for left, and the / key for right (labelled L and R). A 500ms inter-trial interval followed the participant's response, after which the next trial began automatically.

There were 80 trials per emotion (happy, sad, angry), divided equally between congruent (probe behind emotion) and incongruent (probe behind neutral) trials, as well as 40 neutral/baseline (both neutral pictures) trials. As we were only interested in the traditional attentional bias score, we did not use the baseline trials in our analysis. The task was split into two blocks of 140 trials, allowing participants a self-determined break between blocks. A randomization algorithm determined stimulus presentation to ensure that expression, trial type (congruent, incongruent, baseline), and dot location (left, right) occurred with equal frequency in each block, and such that consecutive response presses on either side were limited to 3 to prevent response bias.

Clinical Measures

Difficulties with emotion regulation scale (DERS; Gratz & Roemer, 2004). This 36-item self-report questionnaire consists of six dimensions of emotion regulation: (a) non-acceptance of emotional reactions; (b) difficulty engaging in goal-directed behaviour; (c) difficulty controlling impulses; (d) lack of emotional awareness; (e) limited access to emotion-regulation strategies; and (f) lack of emotional clarity. The DERS begins each item with the phrase, "When I'm upset…". Participants rate each item on a five-point Likert scale from "almost never" (1) to "almost always" (5). Eleven items are reverse scored, and the total score can range from 36 to 180 with a higher score representing increased difficulty with emotion regulation. In non-clinical samples of adults, the DERS has an average total score of 75-80 (Gratz & Roemer, 2004). Only the total score was used in this study. The DERS has high internal consistency (α=.93), good test-retest reliability (ρ_I=.88), and satisfactory construct and predictive validity (Gratz & Roemer, 2004). Within this sample, Cronbach's alpha was .95.

Depression Anxiety and Stress Scale short form (DASS-21; Lovibond & Lovibond, 1996). This self-report questionnaire consists of 21 items divided into three subscales, each

consisting of 7 four-point Likert rating items measuring depression, anxiety and stress. The questionnaire asks the respondent to rate how much each statement applied to them over the past week. The rating scale ranges from 0 (did not apply to me at all) to 3 (applied to me very much, or most of the time). The DASS-21 has high internal consistency and moderate levels of concurrent validity (*r* =.46-.85; Antony et al., 1998). Within the current sample, Cronbach's alpha was .95. The depression subscale was used to measure depressive symptoms. To measure symptoms of anxiety both the anxiety and stress subscales were used, as the anxiety subscale loads onto diagnostic criteria for the various anxiety disorders *except* generalised anxiety disorder (GAD), whereas the stress subscale loads onto diagnostic criteria for GAD (Lovibond & Lovibond, 1995). This allowed us to differentiate between GAD and other types of anxiety. Scores range from 0-28+ for depression (0-4 normal, 5-6 mild, 7-10 moderate, 11-13 severe, 14+ extremely severe), 0-20+ for anxiety (0-3 normal, 4-5 mild, 6-7 moderate, 8-9 severe, 10+ extremely severe), and 0-32+ for stress (0-7 normal, 8-9 mild, 10-12 moderate, 13-16 severe, 17+ extremely severe), with higher scores indicating increased level of symptom severity.

Analysis

Data cleaning and preparation.

Dot Probe task.

Traditional analysis. As per previous research (e.g. MacLeod et al., 1986; Koster et al., 2004) error responses were removed from analysis of RT data. On average, participant accuracy rates were 99.98%. As two participants had error rates at near chance level (52 and 58%) their data were removed from analysis. For each trial type (congruent/incongruent) within each emotion (happy, sad, angry), RTs that were 3 or more standard deviations above each participant's mean RT were removed, as were responses faster than 150ms. This resulted in 1.72% of total responses being removed. For each emotion, an overall bias score for each participant was calculated by taking the mean RT for congruent trials and subtracting it from the mean RT for incongruent trials. Thus, positive scores indicate an attentional bias toward emotional stimuli, whereas negative scores

indicate a bias toward neutral stimuli and away from emotional stimuli. Reliability analyses of the overall bias score identified Spearman-Brown split half reliability to be .52, and Guttman Split-Half .51, both well below acceptable levels of .90.

Fast-dm 30 analysis. Data were cleaned in the same way as for the traditional analysis with the exception that errors were not excluded. The data were coded such that congruent responses were mapped to the upper response threshold and incongruent responses were mapped to the lower threshold, and neutral trials were removed. Fast-dm 30 was run in line with recommendations from Voss, Voss, and Lerche's (2015) tutorial paper. The fast-dm 30 control file was set to allow drift rate and non-decision time to vary based on emotion across participants, and start point and boundary separation were free to vary across participants but not across emotion. Additional parameters of variability in starting point, drift rate, and non-decision time can also be introduced into model analysis. While more parsimonious models can be preferable to more complex models (Lerche & Voss, 2016), it has been recommended to include non-decision time variability in the model in order to achieve stable parameter estimates (Voss et al., 2015; Lerche and Voss, 2016). Therefore, inter-trial variability of non-decision time (s_{t0}) was also allowed to vary across participants. Split half analysis was carried out by taking an odd/even split of the trials from the dot-probe task, and running these two datasets using fast-dm. The resultant parameters from each data set were compared using correlational analysis. Starting point did not correlate between the two halves of the data, however the remaining parameters were significantly related. Boundary separation was highly correlated (r = .89), non-decision time was also highly correlated between halves (angry r = .88, happy r = .89, sad r = .88). However, drift rate was negatively correlated between halves (angry r = -.38, happy r = -.52, sad r = -.49), which indicates minimal consistency between the two halves of data for the drift rate parameter.

Model fit was assessed by the Kolmogorov–Smirnow (KS) statistic provided by the fast-dm 30 software. The KS statistic assesses whether the distribution of observed data differs significantly from a predicted distribution. The KS fit statistic returned by fast-dm 30 did not reveal any

significant deviations between empirical and estimated RT distributions for any of the participants. As Voss et al. (2015) caution against relying solely on the fit statistic returned by fast-dm 30 and recommend plotting the data graphically, comparison of empirical and predicted response time distributions for each expression (happy, angry, sad, neutral) were plotted, along with quantile probability plots comparing empirical and predicted response times for the 25^{th}, 50^{th}, and 75^{th} percentiles of each expression (happy, sad, angry). These plots are displayed in Appendix 1.

DERS and DASS. Clinical measures were checked for skewness and kurtosis, and were within acceptable limits (+/- 2). There were no outliers.

Inferential analysis of cognitive measures. Repeated measures ANOVAs were performed to analyse differences between both response times and parameters. Bonferroni correction was applied, and where sphericity tests indicated it applicable, the Greenhouse Geisser correction was also applied.

Relationships between cognitive and clinical measures. Correlations were carried out to explore whether bias scores or the cognitive components of implicit decision-making were related to scores on the clinical outcome measures of DERS and DASS.

Results

Dot Probe Task

Response times. Correct trial mean RTs for congruent and incongruent trials across the emotions (happy, sad, angry) are plotted in Figure 2.

Attentional Bias. A 2 (trial type: congruent, incongruent) x 3 (emotion: happy, sad, angry) repeated measures ANOVA identified a main effect of emotion, $F(2,180) = 3.35$, $p = .04$, $\eta^2_{partial} = .07$. Pairwise comparisons revealed that happy responses were 4 ms faster than sad responses ($p = .045$), but revealed no significant difference between happy and angry (3 ms, $p = .16$), or angry and sad (1 ms, $p = 1.00$), responses. Analysis showed no main effect of trial type, $F(1,90) = 3.80$, $p = .054$, $\eta^2_{partial} = .04$, nor a significant trial type × emotion interaction, $F(1.84,165.40) = 0.64$, $p = .52$, $\eta^2_{partial} < .007$, indicating that while RTs for happy trials were modestly faster than sad trials

overall, this did not differ significantly by congruence. As such, the data provided no evidence for an attentional bias to any specific emotion.

Figure 2. Dot probe RTs (seconds) by trial type for correct trials. Congruent trials = probe behind emotional face, incongruent trials = probe behind neutral face. Error bars represent within-subjects 95% confidence intervals[2].

Diffusion Model parameters. One-way (emotion: happy, sad, angry) repeated measures ANOVAs revealed no effect of emotion on drift rate (*v*), $F(2, 180) = 2.06$, $p = .13$, $\eta^2_{partial} = .02$, or non-decision time (*t0*), $F(2, 180) = 2.53$, $p = .08$, $\eta^2_{partial} = .03$. One sample *t*-tests indicated that drift rate was positive for all three emotions ($M = 0.15$, $SD = 0.27$, $t(90)=5.32$, $p<.001$, $d=0.56$ for happy; $M=0.07$, $SD=0.29$ $t(90)=2.39$, $p=.019$, $d=0.24$ for sad; $M=0.13$, $SD=0.27$, $t(90)=4.43$, $p<.001$, $d=0.48$ for angry), indicating a slightly faster drift rate for congruent than for incongruent trials. However, evidence did not accumulate more quickly for one emotion over another. The ANOVA identified that there was also no difference in non-decision time across emotions ($M=0.30$, $SD=0.04$

[2] All confidence intervals throughout this paper were calculated using the Cousineau-Morey method (Morey, 2008).

for happy; *M*=0.31, *SD*=0.04 for sad; *M*=0.13, *SD*=0.04 for angry), indicating no evidence of disorder congruent non-decisional processes – i.e., motor slowing or hyperactivity across the sample.

Clinical measures. Means and standard deviations for the clinical measures are shown in Table 2. Mean scores on the clinical measures are at the lower end of the clinical range; however, the standard deviations indicate that there is a reasonable amount of variance within the sample, indicating a reasonable distribution of clinically significant scores.

Correlations (Table 3) revealed that attentional bias toward happy faces and attentional bias toward angry faces were correlated with emotion regulation scores, such that stronger attentional biases toward happy and angry emotions were associated with better emotion regulation. In addition, angry attentional bias scores were associated with lower scores of depression and stress (GAD).

None of the attentional bias scores correlated with scores on the anxiety subscale (other anxiety disorders). These results do not support the hypothesis that biases toward negative emotional stimuli are associated with increased emotional regulation difficulty. Instead they suggest that avoidance of emotion, rather than a focus on emotion, is more symptomatic of emotional difficulties. It is important to note that these correlational values are prior to Bonferroni correction. Whilst in the expected direction, once correction is applied, results become non-significant, so we must interpret these results with caution.

Boundary separation was negatively correlated with scores on the DERS, such that a more conservative decision style was associated with better emotion regulation. Drift rate for angry faces was correlated with depression and stress scores, such that a higher drift rate for congruent stimuli (positive drift) in the angry condition was linked to higher depression and stress scores. This suggests that individuals whose attention was captured by angry stimuli tended to have higher clinical scores. This is in the opposite direction to the findings of the RT analysis.

Table 2.

Clinical Measures

	N	Minimum	Maximum	Mean	Std. Deviation
DERS Total	91	47	155	91.76	23.52
Depression	91	0	40	10.46	9.92
Anxiety	91	0	34	7.65	7.85
Stress	91	0	38	14.13	9.17

Table 3

Correlations between cognitive and clinical measures (N=91)

		Depression	Anxiety	Stress	DERS
Response Times					
Attentional Bias	happy	-.123	-.120	-.153	-.246*
	sad	.035	.055	.093	.021
	angry	-.277**	-.105	-.216*	-.255*
Fast-dm 30 parameters					
Boundary separation		-.163	-.072	-.152	-.237*
Drift rate	happy	.015	.137	.097	.145
	sad	-.056	-.076	-.031	-.027
	angry	.282**	.199	.233*	.195
Non-decision time	happy	-.09	-.127	-.122	-.305**
	sad	-.081	-.155	-.153	-.282**

	angry	-.059	-.127	-.121	-.231*

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Discussion

The aim of this paper was to assess whether the diffusion decision model, as applied via fast-dm 30, provides a more sensitive and reliable way to measure attentional bias in the dot probe task than the bias score derived from correct mean RTs. Additionally, we were interested in whether the diffusion model parameters could offer greater insight into the relationship between cognitive processes that underlie attentional bias and clinical symptoms than the traditional bias score. To this end, we explored the relationships between the parameter values returned by fast-dm 30 and scores on the DERS and DASS.

Traditional RT analysis of the dot probe task did not reveal an overall attentional bias for happy, sad, or angry emotional expressions. Given the non-clinical nature of the sample, this overall effect is perhaps not surprising. However, correlational analyses suggested that attentional biases for happy and angry faces may be associated with better emotion regulation, and attentional bias for angry faces may be linked to lower scores on the stress (GAD) and depression scales. Additionally, attentional bias for sad faces did not correlate with scores on depression. These trends are contrary to theoretical expectations that biased attention *toward* emotional stimuli is a risk factor for psychopathology (Beck, 1976), and instead suggest that the ability to attend to emotional information may be a protective factor. In this sample, it seems that individuals who have difficulty regulating emotion avoid emotion rather than seek it out.

Analysis of diffusion model parameters for the dot probe task revealed that drift rate was slightly positive for all conditions, indicating a tendency to engage with emotional faces more strongly than neutral faces. While this result was significant, it is important to note that this was a small to medium effect (Cohen's d ranging from 0.24 -0.56). Additionally, evidence did not accumulate more quickly for one emotion over another, and there was no difference in non-decision times for any emotion. This mirrors the lack of attentional bias captured through traditional RT

analysis, and may reflect the recruitment of a non-clinical sample. However, drift rate was linked to depression and stress scores such that a faster rate of evidence accumulation for angry faces was indicative of higher levels of self-reported depression and generalised anxiety. These results are in line with Beck's content specificity hypothesis (1976) to the extent that evidence for angry information accumulated more quickly for those with higher levels of generalised anxiety. According to the content specificity hypothesis, however, depression should be associated with increased evidence accumulation of sad faces, which was not the case here. It may be that the current sample did not have a high enough proportion of individuals with clinical levels of depression to capture the relationship, or it may be that depression is not as strongly associated with an attentional bias toward sad information. To this end, the link between attentional bias for sad stimuli and depression has typically been found in studies that use stimulus presentation times of 1000ms (Bradley, Mogg, & Lee, 1997; Gotlib et al., 2004). Some researchers suggest that attentional bias for sad stimuli is more strongly related to a difficulty disengaging attention rather than attentional capture, which is argued to account for the longer stimulus presentation times required to observe an attentional bias for sad stimuli (e.g., Joorman, 2010). If the relationship between depression and attention is more strongly related to disengagement difficulties, drift rate would not identify this relationship at the present stimulus presentation time of 500ms.

While neither boundary separation nor non-decision time for any emotional expression were correlated with anxiety or depression specifically, both boundary separation and non-decision time for all emotions were negatively correlated with emotion regulation scores.[3] These results imply that a more conservative decision style is linked with better emotion regulation, and faster non-decisional processes are related to an increased difficulty regulating emotion. Theoretically, these findings are in line with Eysenck et al.'s (2007) attentional control theory, which states that cognitive control is inhibited during a hyper-vigilant response to emotion and as such, the ability to regulate responses to emotion is hindered.

[3] Importantly, scores on the DASS were significantly correlated with scores on the DERS (r= .64, .54, .65 for depression, anxiety and stress subscales, respectively).

Taken altogether, our results suggest that the traditional bias score and the results of the diffusion model are conflicting in their relationship with clinical measures. The relationship between mean RTs and clinical measures suggests that an avoidance of emotion, rather than a focus on emotion, is more indicative of emotional difficulties. The parameters of the diffusion model, however, were linked to clinical measures such that a faster accumulation of evidence for angry stimuli *was* linked with depression and GAD, while slower non-decisional processes and more cautious responding were a protective factor to emotion regulation difficulties. The only theoretically plausible explanation that can account for these conflicting results is the notion of a suppression effect as identified by the separation of the diffusion parameters.

The results from the diffusion decision model are in line with Beck's (1976) content specificity hypothesis for anxiety, and support Eysenck et al.'s (2007) attentional control theory. While the results do not support the content specificity hypothesis for depression in that sad expressions did not capture attention as expected, capture by angry expressions was linked to *both* anxiety and depression, suggesting a broader connection between negative emotion and emotional disorders. Based on these findings, the fast-dm 30 parameters seem to be a more promising account of the data than traditional bias scores. However, the internal consistency of the drift rate parameter in the current sample was inadequate.

There are several possible reasons that the drift rate parameter demonstrated low reliability. It may be that the way the data were mapped using fast-dm does not offer internal stability of the drift rate parameter. Alternatively, it may be that attentional bias is not a function of drift-rate. Or it may be that the low reliability found in both the traditional attentional bias score, and the drift rate parameter is simply due to the absence of an attentional bias in the current sample. This is the first time, to our knowledge, that the diffusion decision model has been applied to data from the dot-probe task. Therefore, replication of these results is needed to confirm the fit of the model, as well as the pattern of results obtained, before we can have high confidence in the conclusions. In the interim, we can tentatively conclude that the acceptable fit statistics of the diffusion model, and the

results supporting current theoretical understanding, offer more insight into the cognitive processes underlying attentional bias as identified by the diffusion model, than RT analysis alone.

As more attentional bias researchers apply the diffusion model to their existing data, we will have a wider evidence base to draw from, and we can begin to draw firmer conclusions as to the efficacy of the model as applied to dot probe data. If our tentative findings are replicated, and thus, deemed reliable, then the diffusion model can be considered a more comprehensive form of analysis. We will be able to confidently identify the cognitive processes responsible for attentional bias and offer more insight into the cognitive processes that underlie attentional bias to affective stimuli and its link to psychopathological symptoms. If the suppression effect holds up, it may account for the conflicting results that have been seen in attentional bias modification research. Specifically, there may be changes in some of the underlying cognitive processes that cannot be detected by traditional bias score analysis alone.

Applying the diffusion model to attentional bias modification data has implications for clinical research. Specifically, such application will enable us to identify *how* successful bias modification works, i.e., which parameters shift from pre-training to post training. Additionally, by applying the model to data from the control group of a modification study (pre- and post-training measures), test-retest reliability of the fast-dm 30 parameters could be assessed, increasing confidence in the model's suitability as a method of analysis for the dot probe task. From there, we could use the model to identify suitable candidates for attentional bias modification treatment. Further, we can devise paradigms that target one or more of the specific cognitive processes that are determined to be responsible for attentional biases with the aim of improved treatment outcomes. For example, the parameters of the diffusion model tentatively confirm that the bias found in raw RTs for depression does not seem to be occurring during the first 500ms of decision processing. Therefore, by exploring when it is best to target a bias, i.e., how long to present stimuli for, could enhance the efficacy of training. Additionally, the current results suggest that more cautious responding, which is related to better cognitive control, may mitigate depression and anxiety

onset/severity by its link with better emotion regulation. As such, tasks that strengthen cognitive control may provide another beneficial treatment approach.

Conclusion

The aim of this research was to explore whether Ratcliff's diffusion model, as applied using fast-dm 30, could provide reliable, and more comprehensive information, than the current attentional bias measure of mean RTs. Our results, while exploratory, do indicate acceptable fitting of the data as well as trends that are in line with current theory. Specifically, links between fast-dm 30 parameters and clinical measures were in line with both the content specificity hypothesis and attentional control theory. These findings indicate that cognitive control may be a protective factor for emotion regulation difficulties, and in turn, anxiety and depression. Importantly, our results suggest that the application of the diffusion decision model to attentional bias is an area that warrants further research. In so doing, application of the diffusion model to existing dot probe data could provide extensive exploratory *and* confirmatory analysis.

References

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), 1–24. https://doi.org/10.1037/0033-2909.133.1.1

Beard, C., Sawyer, A. T., & Hofmann, S. G. (2012). Efficacy of Attention Bias Modification Using Threat and Appetitive Stimuli: A Meta-Analytic Review. *Behavior Therapy, 43*(4), 724–740. https://doi.org/http://dx.doi.org/10.1016/j.beth.2012.01.002

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. International Universities Press.

Bradley, B. P., Mogg, K., & Lee, S. C. (1997). Attentional biases for negative information in induced and naturally occurring dysphoria. *Behaviour Research and Therapy, 35*(10), 911–927.

Bradley, B. P., Mogg, K., & Millar, N. H. (2000). Covert and overt orienting of attention to emotional faces in anxiety. *Cognition & Emotion, 14*(6), 789–808.

Cisler, J. M., Bacon, A. K., & Williams, N. L. (2009). Phenomenological Characteristics of Attentional Biases Towards Threat: A Critical Review. *Cognitive Therapy Research, 33*(2), 221–234. https://doi.org/10.1007/s10608-007-9161-y

Cisler, J. M., & Olatunji, B. O. (2012). Emotion Regulation and Anxiety Disorders. *Current Psychiatry Reports, 14*(3), 182–187. https://doi.org/10.1007/s11920-012-0262-2

Ekman, P., & Friesen, W. V. (1986). A new pan-cultural facial expression of emotion. *Motivation and Emotion, 10*(2), 159–168. https://doi.org/10.1007/BF00992253

Etkin, A., Büchel, C., & Gross, J. J. (2015). The neural bases of emotion regulation. *Nature Reviews Neuroscience, 16*(11), 693–700. https://doi.org/10.1038/nrn4044

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion (Washington, D.C.), 7*(2), 336–353. https://doi.org/10.1037/1528-3542.7.2.336

Gotlib, I. H., Krasnoperova, E., Yue, D. N., & Joormann, J. (2004). Attentional biases for negative interpersonal stimuli in clinical depression. *Journal of Abnormal Psychology*, *113*(1), 127.

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment*, *26*(1), 41–54.

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*(11), 982–990. https://doi.org/http://dx.doi.org/10.1016/j.biopsych.2010.07.021

Joormann, J., & Quinn, M. E. (2014). Cognitive processes and emotion regulation in depression. *Depress Anxiety*, *31*(4), 308–15. https://doi.org/10.1002/da.22264

Koster, E. H., Crombez, G., Verschuere, B., & De Houwer, J. (2004). Selective attention to threat in the dot probe paradigm: Differentiating vigilance and difficulty to disengage. *Behaviour Research and Therapy*, *42*(10), 1183–1192.

Koster, E. H., Verschuere, B., Crombez, G., & Van Damme, S. (2005). Time-course of attention for threatening pictures in high and low trait anxiety. *Behaviour Research and Therapy*, *43*(8), 1087–1098.

Lerche, V., & Voss, A. (2016). Model Complexity in Diffusion Modeling: Benefits of Making the Model More Parsimonious. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01324

Lovibond, S., & Lovibond, P. F. (1996). *Manual for the depression anxiety stress scales*. Psychology Foundation of Australia.

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, *95*(1), 15.

Mathews, A., & Mackintosh, B. (1998). A Cognitive Model of Selective Processing in Anxiety. *Cognitive Therapy and Research*, *22*(6), 539–560. https://doi.org/10.1023/A:1018738019346

Mathews, A., & MacLeod, C. (1985). Selective processing of threat cues in anxiety states. *Behaviour Research and Therapy*, *23*(5), 563–569.

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

Mogg, K., & Bradley, B. P. (2005). Attentional bias in generalized anxiety disorder versus depressive disorder. *Cognitive Therapy and Research*, *29*(1), 29–45.

Peckham, A. D., McHugh, R. K., & Otto, M. W. (2010). A meta-analysis of the magnitude of biased attention in depression. *Depression and Anxiety*, *27*(12), 1135–1142.

Pergamin-Hight, L., Naim, R., Bakermans-Kranenburg, M. J., van, Ij. M. H., & Bar-Haim, Y. (2015). Content specificity of attention bias to threat in anxiety disorders: a meta-analysis. *Clinical Psychology Review*, *35*, 10–8. https://doi.org/10.1016/j.cpr.2014.10.005

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25.

Price, R. B., Kuckertz, J. M., Siegle, G. J., Ladouceur, C. D., Silk, J. S., Ryan, N. D., … Amir, N. (2015). Empirical recommendations for improving the stability of the dot-probe task in clinical research. *Psychological Assessment*, *27*(2), 365–76. https://doi.org/10.1037/pas0000036

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.

Ratcliff, R., & Smith, P. (2015). Modeling simple decisions and applications using a diffusion model. In *The Oxford handbook of computational and mathematical psychology*. Retrieved from

http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199957996.001.0001/oxfordhb-9780199957996-e-3?mediaType=Article

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*(2), 333.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438–481.

Rodebaugh, T. L., Scullin, R. B., Langer, J. K., Dixon, D. J., Huppert, J. D., Bernstein, A., … Lenze, E. J. (2016). Unreliability as a threat to understanding psychopathology: The cautionary tale of attentional bias. *Journal of Abnormal Psychology*, *125*(6), 840.

Sakagami, M., & Pan, X. (2007). Functional role of the ventrolateral prefrontal cortex in decision making. *Current Opinion in Neurobiology*, *17*(2), 228–233. https://doi.org/10.1016/j.conb.2007.02.008

Schmukle, S. C. (2005). Unreliability of the dot probe task. *European Journal of Personality*, *19*(7), 595–605.

Shechner, T., Britton, J. C., Pérez-Edgar, K., Bar-Haim, Y., Ernst, M., Fox, N. A., … Pine, D. S. (2012). Attention biases, anxiety, and development: toward or away from threats or rewards? *Depression and Anxiety*, *29*(4), 282–294.

Snyder, H. R. (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: A meta-analysis and review. *Psychological Bulletin*, *139*(1), 81–132. https://doi.org/10.1037/a0028727

Sobin, C., & Sackeim, H. A. (1997). Psychomotor symptoms of depression. *The American Journal of Psychiatry*, *154*(1), 4.

Staugaard, S. R. (2009). Reliability of two versions of the dot-probe task using photographic faces. *Psychology Science Quarterly*, *51*(3), 339–350.

Švegar, D., Kardum, I., & Polič, M. (2013). Happy Face Superiority Effect in Change Detection Paradigm. *Psychological Topics*, *22*(2), 249–269.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., … Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

Tovote, P., Fadok, J. P., & Lüthi, A. (2015). Neuronal circuits for fear and anxiety. *Nature Reviews Neuroscience*, *16*(6), 317–331. https://doi.org/10.1038/nrn3945

Vandekerckhove, J., & Tuerlinckx, F. (2008). Diffusion model analysis with MATLAB: A DMAT primer. *Behavior Research Methods*, *40*(1), 61–72. https://doi.org/10.3758/BRM.40.1.61

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, *32*(7), 1206–1220. https://doi.org/10.3758/BF03196893

Voss, A., & Voss, J. (2007). Fast-dm: a free program for efficient diffusion model analysis. *Behavior Research Methods*, *39*(4), 767–775.

Voss, A., Voss, J., & Lerche, V. (2015). Assessing cognitive processes with diffusion model analyses: a tutorial based on fast-dm-30. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00336

Wagenmakers, E.-J., van der Maas, H. L. J., & Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*(1), 3–22.

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Using diffusion models to understand clinical disorders. *Journal of Mathematical Psychology*, *54*(1), 39–52. https://doi.org/http://dx.doi.org/10.1016/j.jmp.2010.01.004

Williams, J. M., Watts, F. N., MacLeod, C., & Mathews, A. (1997). *Cognitive psychology and the emotional disorders. (2nd ed.).*

Zvielli, A., Bernstein, A., & Koster, E. H. W. (2015). Temporal Dynamics of Attentional Bias. *Clinical Psychological Science*, *3*(5), 772–788. https://doi.org/10.1177/2167702614551572

Chapter 4: Study 2

The impact of stimulus strength on capturing affective interpretation bias and emotion regulation as measured by the yes/no task: A diffusion model analysis

Amy Manuel, Eva Kemps, Jason S. McCarley

School of Psychology, Flinders University, Adelaide, Australia

Author Note

Jason S McCarley, School of Psychology, Flinders University, is now at the College of Liberal Arts, Oregon State University, USA

Correspondence concerning this article should be addressed to Amy Manuel, School of Psychology, Flinders University, GPO Box 2100, Adelaide, SA 5001, Australia. Electronic mail may be sent to Amy.Manuel@flinders.edu.au

Abstract

Efforts to modify cognitive biases in anxiety and depression to alleviate symptoms is a burgeoning area of research. Knowing which cognitive processes to target will help with treatment outcomes. Our aim was to use drift diffusion analysis to identify which cognitive processes are affected by interpretation of ambiguous stimuli. Participants (N = 93) performed a yes/no task to classify, and thus interpret, whether a presented facial expression belonged to the target category. One group of participants were presented with intense facial expressions (*n* = 48), and another group were presented with a mix of both intense and milder, more ambiguous facial expressions (*n* = 45). Results indicate that both perceptual bias (drift rate) and decisional bias (starting point) were affected by the inclusion of ambiguous stimuli. The participants presented with the ambiguous stimuli had lower drift rates and reduced starting point bias than the participants presented with the non-ambiguous stimuli. Thus, perceptual bias *and* decisional biases are affected by interpreting ambiguous expressions. Application in a clinical sample may offer insight into the processes responsible for interpretation bias in anxiety and depression. This knowledge would aid in devising and applying more targeted bias modification treatments, in efforts to alleviate symptoms of these disorders.

Key words: Interpretation bias, Diffusion decision model, Emotion regulation difficulties, Yes/No task.

The impact of stimulus strength on capturing affective interpretation and emotion regulation as measured by the yes/no task: A diffusion model analysis

An implicit tendency to interpret benign, ambiguous, or mild affective information in a heightened emotional manner is known as interpretation bias (Beck, 1976). Anxiety has been associated with a tendency to interpret ambiguous information with an inflated perception of threat (e.g. Eysenck, Mogg, May, Richards, & Mathews, 1991). In the same way, depression has been associated with a stronger implicit dysphoric interpretation of ambiguous information (e.g., Gotlib & Joormann, 2010). These interpretation biases are linked with increased difficulty regulating strong emotion, which in turn, intensifies symptoms of anxiety and depression (Cisler & Olatunji, 2012; Joormann & Quinn, 2014). Consequently, interpretation bias and difficulty regulating emotion are theorized to increase vulnerability to anxiety and depression (Cisler & Olatunji, 2012; Joormann & Quinn, 2014).

Emotion regulation is the ability to monitor, evaluate, and modify the magnitude and duration of one's emotional reactions in order to achieve desired personal goals (Svaldi et al., 2012). The most common emotion regulation goals in everyday life are to decrease negative emotion and increase positive emotion (Gross, 2013). Emotion regulation strategies incorporate intrinsic (self-regulated) and extrinsic (regulated by others) processes that provide awareness, understanding and acceptance of our emotional states. Individuals learn to regulate their emotions during infant development through extrinsic processes (Gross & Thompson, 2007), i.e., co-regulation from parents and caregivers. A well-regulated emotional system enables an individual to act in the desired manner regardless of emotional state (Svaldi et al., 2012). Emotion regulation processes can be explicit, e.g. consciously trying to calm one's nerves before a performance, or implicit, e.g. unconsciously diverting one's gaze from threatening stimuli (Gross, 2013).

The ability to regulate strong emotion, either implicitly or explicitly, is an important part of social functioning and good mental health (Gross, 2013). Reacting impulsively to emotion is detrimental to social connectedness, as it can alienate others and increase feelings of isolation, and

consequently, increase negative emotion (Gross, 2013). As a result, emotion regulation difficulties are transdiagnostic in nature, that is, they exist across many psychological diagnoses and underlie a number of emotional disorders (Aldao, Nolen-Hoeksema, & Schweizer, 2010). Specifically, the inability to self-regulate is associated with experiences of increased negative emotion, which can exacerbate disorder-relevant symptoms, leading to the onset and maintenance of affective disorders such as anxiety and depression (Gross, 2013).

Emotion Generation and Emotion Regulation

Gross and Thompson (2007) developed the *process model of emotion regulation* (Figure 1), an evidence processing model that treats each step in the generation of emotion as a potential target for regulation. In the model, emotion generation starts with a situation which the individual perceives as relevant to their goals. The individual pays attention to the situation, and because it has specific meaning to them, it elicits a multi-system (physiological, behavioural, and subjective experiential) response. In other words, the appraisals the individual makes while attending to the situation result in an emotional response. That response then changes the individual's perception of the situation, which results in a new process of evaluation and the potential to generate new emotions, creating a cycle of emotion generation and response.

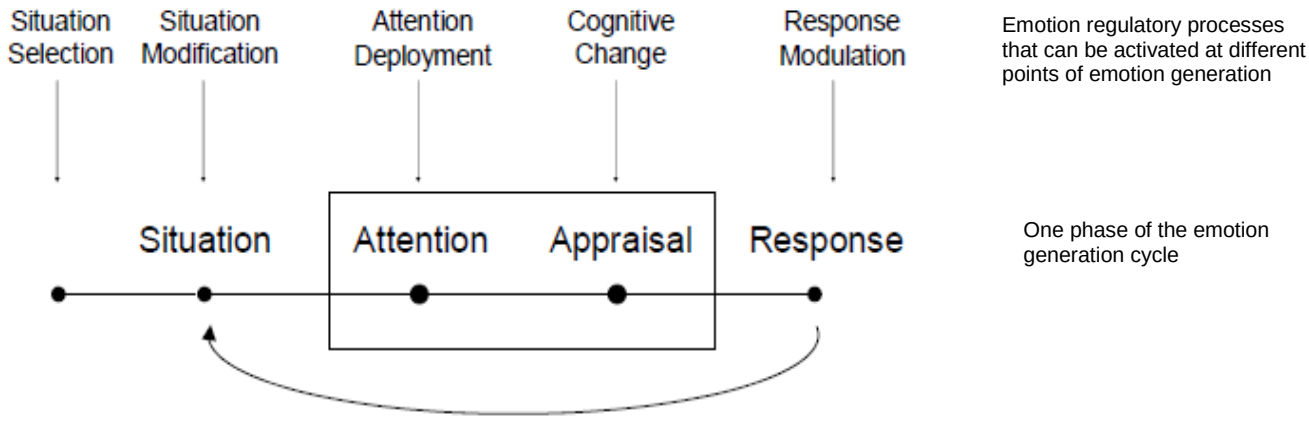


Figure 1. Process model of emotion regulation (Gross & Thompson, 2007)

During a single phase of the emotion generation cycle there are five points at which regulation can occur: situation selection, situation modification, attentional deployment, cognitive change, and response modulation. At each of these points a specific process can be employed to intrinsically or extrinsically regulate emotion. This regulation attempts to dampen, intensify, or maintain the emotion, depending on the individual's goal. Emotion regulation strategies employed during each of these processes have a different consequence, and each of these consequences will change the course of the individual's response to the situation, in turn modifying the emotional experience.

Generally, individuals regulate their emotions using a number of strategies concurrently along different points of the emotional processing continuum (Gross & Thompson, 2007). However, in some individuals, aberrant deployment of attention toward emotional stimuli, difficulty disengaging from such stimuli, or mistakes processing them can generate emotion that is overwhelming for the individual. It is these biased cognitive processes that are hypothesised to underlie attentional and interpretation biases. Such biases impact the regulation processes of attention deployment and cognitive change, fostering dysfunctional emotion regulation strategies that increase difficulty coping with emotion rather than decrease it. When faced with overwhelming emotion and a lack of functional emotion regulation skills, it can be very difficult to regulate emotion effectively, and it is this combination of dysfunctional emotional processing and lack of regulatory skills that is thought to be a causal process in a number of psychological disorders such as anxiety and depression (Browning, Holmes, & Harmer, 2010).

Attention deployment entails looking away from something to alleviate the emotional response it arouses (distraction) or focusing attention towards something to increase or maintain the emotional response it arouses (concentration). Concentration on stimuli that evoke negative emotions is referred to as rumination, and is known to increase depressive symptoms (Gross & Thompson, 2007). In someone with an adaptive regulatory process, the initial response to negative

information is to avoid it. However, for someone for whom this process has been disrupted, it can lead to maladaptive behaviours such as fixation and rumination. As maladaptive strategies are often implicit, the awareness of, and ability to regulate them is usually not within the individual's control, and as such increases the negative emotion the individual experiences (Gross & Thompson, 2007).

Cognitive change refers to the process of reappraisal. Reappraising a situation can alter its emotional impact, making it more, or less, significant. For example, if an individual with a negative interpretation bias has an ambiguous or benign interaction with someone they would prefer to avoid, their initial appraisal of the situation will be negative. Instead of reappraising the situation in a positive way (i.e., "Okay, that wasn't so bad."), they may continue to interpret the situation negatively (i.e., "That was horrible!"). This maladaptive interpretation works to increase negative emotion rather than to alleviate it, and may result in overwhelming emotion that requires response modulation.

Interpretation bias has its strongest impact in the attention and appraisal stages of emotion generation. The *content specificity hypothesis* suggests that individuals are drawn to information that reflects their current mood (Beck, 1976). Therefore, when an individual with a negative interpretation bias is in a negative mood, they will interpret environmental stimuli as negative, and be drawn to this information. *Attentional control theory* (Eysenck & Calvo, 1992) posits that attention shifting difficulties can result in fixating on self-referent stimuli. Thus, an inability to shift attention from self-referent information can result in fixation and rumination. Further, if an individual's cognitive control is inhibited, their ability to regulate their emotional response is also inhibited. As a result, fixation and rumination without cognitive control result in difficulty regulating strong emotion (Joormann & Siemer, 2011) and cause a heightened emotional response, i.e., exacerbated symptomology. Hence, a bias toward negative information can impede successful emotion regulation.

Measurement and Analysis of Interpretation Bias and Cognitive Control

Interpretation bias is typically measured using ambiguous written scenarios. Participants are presented with a series of ambiguous phrases or scenarios with a final word presented as a fragment that they must solve. The valence of the ambiguous scenario is determined by this final word. For example, "You've just started reading a new book that you bought and you find it to be …", with results being scored as negative (e.g. "boring") or positive (e.g. "interesting") (Hallion & Ruscio, 2011). These scores for positive and negative interpretations are then analysed to examine differences between clinical groups, and while they identify group differences, they do not offer insight into the cognitive processes that underlie the decision making.

Cognitive control is primarily measured using computer tasks that evoke a Stroop effect or Simon effect. These tasks employ a stimulus-stimulus (Stroop effect) or stimulus-response (Simon effect) conflict that requires cognitive control to inhibit an automatic behaviour and execute the response correctly (O'Leary & Barber, 1993). These tasks are typically timed tasks from which response times for differing stimuli, for example affective versus neutral, are compared between groups. When analysed comparing response time alone, these tasks do not allow for inferences to be made about the cognitive processes that underlie the decision making.

Thus, the aim of the current study was to examine the cognitive processes that underlie interpretation bias and cognitive control, and their relationship to emotion regulation using an evidence accumulation model of analysis. Evidence accumulation models can tease apart the cognitive processes that are responsible for interpretation bias and cognitive control. These models separate the underlying cognitive processes in speeded binary decision tasks using a small number of parameters. The parameters of the models represent well-defined and theoretically meaningful psychological constructs, and offer inferences into the cognitive processes that are hidden from traditional analysis (Donkin, Brown & Heathcote, 2011; Voss, Rothermund & Voss, 2004; Voss & Voss, 2007). By isolating and quantifying the components of cognition that underlie speeded

decision making, evidence accumulation models can provide evidence about cognitive processes that are hidden from traditional analysis (Voss, Rothermund, & Voss, 2004; Voss & Voss, 2007).

However, as evidence accumulation models can only map data from speeded binary decision tasks, the data from popular measures of interpretation bias are not suited to this kind of analysis. Therefore, we have chosen to use an alternative, two choice speeded response task to measure interpretation bias in this study. By using a yes/no task based on an affective go/no-go paradigm (Tottenham, Hare, & Casey, 2011) we were able to measure both interpretation bias and cognitive control using the same task.

The go/no-go association task (GNAT; Nosek & Banaji, 2001) uses mean RTs and signal detection measures to evaluate automatic preferences to stimuli that may reveal individual differences in attitude or interpretation. Traditionally used to discern social bias toward in-groups and out-groups, the GNAT was designed to assess judgments that reflect automatically activated evaluation without the participant's conscious awareness or control (Nosek & Banaji, 2001). More recently the GNAT has also been used within clinical research to assess the role of cognitive control in attentional processes (e.g. Erickson et al., 2005; Gole et al., 2012), as well as the ability to discriminate mood congruent emotional stimuli from mood incongruent stimuli (e.g. Tottenham et al., 2010).

The GNAT requires participants to identify a stimulus that is presented on the screen as belonging to a target category (go) or not (no-go). Stimuli that belong to the category in question are termed targets, and stimuli that do not belong are termed distractors. Targets are presented at a higher rate than distractors to instigate a propensity for "go" responses, increasing the cognitive effort needed to inhibit responses to distractor stimuli. This increased cognitive load on executive functioning is hypothesised to tax an individual's inhibitory control (Schulz et al., 2007), that is, their ability to inhibit the "go" response that is expected. When the target is presented, participants press a key, and when a distractor is presented, they make no overt response. The yes-no version of the GNAT is very similar except that response keys are labelled "yes" and "no", and participants are

required to respond to the distractor stimuli by pressing a different button to identify them. When the stimulus presented is strongly associated with an attribute in line with the participant's own implicit beliefs, response times will be faster, indicating the presence of an interpretation bias.

Interpretation bias is primarily identified through judgments of ambiguous stimuli. Emotional expressions of stronger intensity are easier to identify than ambiguous expressions. Thus, the latter result in longer response times (RTs) and increased error rates (Palmer, Huk, & Shadlen, 2005). As such, if strongly emotive stimuli are used, a ceiling effect can mask differences between participants of different clinical status. In contrast, when the stimuli are more ambiguous, individuals with an interpretation bias will identify the ambiguous stimuli as emotive more quickly than those without an interpretation bias. This is hypothesised to be a function of inhibited cognitive control, which is associated with impulsive responding and decision making, the same processes hypothesised to underlie emotion regulation difficulties (Joormann & Quinn, 2014).

To obtain measures of cognitive bias from the GNAT family of tasks, some researchers use mean response times, and omission and commission errors (e.g. Erickson et al., 2005; Gole et al., 2012; Waters & Valvoi, 2008). Others use signal detection analysis (Nosek and Banaji, 2001; Pacheco-Unguetti et al., 2012; Redick et al., 2011; Schulz et al., 2007; Tottenham et al., 2011), which separates perceptual bias from decisional bias. Signal Detection theory categorizes D prime (d') as a measure of perceptual bias that captures an individual's sensitivity to the stimuli (Nosek and Banaji, 2001; Pacheco-Unguetti et al., 2012; Redick et al., 2011; Schulz et al., 2007), and response caution (C) as a measure of decisional bias that identifies whether individuals are more conservative or liberal in their responding (Lynn & Barrett, 2014). Perceptual bias is a function of the properties of the stimulus, such as physical quality, or semantic clarity, and thus is dependent on elements beyond an individual's control. Whereas, decisional bias is based on an individual's inclination to identify a stimulus as belonging to the target category, which is often based on experience and expectation. In this study, we used a different method of analysis derived from evidence accumulation models that can separate cognitive processes in a similar, but more specific

manner than Signal Detection theory, by also considering the time course of the decision process (White et al., 2016).

We have chosen to use the *diffusion decision model* (Ratcliff, 1978; Ratcliff & McKoon, 2008; Ratcliff & Smith, 2015) because it is one of the most broadly applied of the evidence accumulation models. The diffusion decision model has been shown to account well for the data from two-choice yes-no tasks (Ratcliff & McKoon, 2008), and the two-choice version of the go/no go task can be more easily mapped with the diffusion model than the traditional go/no go version (Gomez, Ratcliff, & Perea, 2007). Importantly, the addition of an inhibitory response button (i.e., the "no" button as opposed to a timed-out non-response) does not affect the integrity of the cognitive processes being measured (Shanoy & Angela, 2012). The interpretational validity of the diffusion decision model has been confirmed (Voss, Rothermund & Voss, 2004), as has the ability of the model to robustly detect moderate changes to the response distribution across trials (Ratcliff, 2013). Ratcliff's diffusion model fits behavioural data as well as, or better than, competing models (Ratcliff & Smith, 2015, 2004; White et al., 2010).

Evidence accumulation models of analysis can be difficult to fit to data, and as such, much of the work analysing clinical disorders using evidence accumulation models has come from researchers with a strong mathematical psychology background (e.g. Heathcote et al., 2015; White, Ratcliff, Vasey, & McKoon, 2010; White, Skokin, Carlos, & Weaver, 2016). However, with several methods of fitting the models to data now available to researchers (Fast-dm, Voss & Voss, 2007; EZ diffusion, Wagenmakers, van der Maas, & Grasman, 2007; DMAT, Vandekerckhove & Tuerlinckx, 2008), the models are becoming easier to use and apply for researchers without a mathematical psychology background.

This study applied the diffusion decision model using the program *fast-dm (*Voss & Voss, 2007), to measurement data from the yes-no task to explore the cognitive processes that are responsible for interpretation bias in response to facial expressions, and emotion regulation difficulties. We have chosen to use facial stimuli as opposed to written words due to the salience

and universality of facial expressions (Bradley, Mogg, & Millar, 2000; Ekman & Friesen, 1986). In line with Beck's (1976) content specificity hypothesis linking attentional processes to mood, the link between emotional disorders and threatening and dysphoric information, we have chosen to use happy, sad, and angry expressions. By comparing judgments of non-ambiguous expressions with judgments of mild, and thus ambiguous expressions, we aimed to examine which cognitive processes are affected by ambiguous information in a non-clinical sample, i.e., which processes are responsible for interpretation bias.

In identifying the processes that underlie interpretation bias, we can better understand how we may be able to more effectively target treatments to minimise the negative impact of interpretation bias in everyday life. Further, we explored the concurrent validity of cognitive control as a measure of emotion regulation difficulty. To date, research into emotion regulation difficulties has commonly used self-report measures such as the Difficulties with Emotion Regulation Scale (DERS; Gratz & Roemer, 2004a). While the DERS itself has good psychometric properties, the inability to remove response bias for socially acceptable answers, and a lack of self-awareness of emotions in some populations (i.e. alexithymia), are a concern with all self-report measures. Therefore, the benefit of a cognitive measure that can capture emotion regulation difficulties irrespective of an individual's desire to respond in a socially acceptable manner or their lack of awareness of their emotions offers a potential increase in construct validity.

The Diffusion Decision Model

The diffusion decision model is based on the theory that when an individual chooses between two options, they accumulate evidence for one or the other option until a decision threshold is reached, and a response is executed. The model decomposes behavioural data - accuracy, mean response times, and response time distributions – obtained from speeded two-choice decision making performance into four parameters: boundary separation, relative starting point, drift rate and non-decision time (see Figure 2 and Table 1).

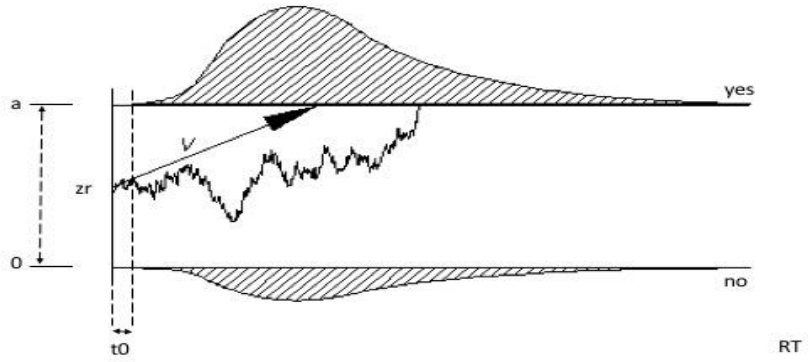


Figure 2. Diffusion decision model response parameters, boundary separation (*a*), relative starting point (z_r), drift rate (*v*), non-decision time (*t0*) for a decision where the response alternatives are yes and no, as per the yes-no task. (Diagram adapted from Voss & Voss, 2007; used with permission)

Boundary separation (*a*). The participant's response caution is reflected in the separation between a lower decision boundary (0) and an upper decision boundary (*a*), each of which corresponds to one of the two available choices, "yes" and "no". Because the value of the lower boundary is fixed at 0, a larger value of *a* indicates a larger boundary separation, meaning that the individual requires more evidence to reach a decision (Voss & Voss, 2007). A larger boundary separation, holding all else equal, will result in responses that are on average slower but more accurate. Boundary separation is our measure of cognitive control.

Relative starting point (z_r). Relative starting point is the point between the decision boundaries from which evidence accumulation starts, measured as a proportion of boundary separation. This parameter reflects the bias an individual has for one response option over another based on their expectancy of that response derived from past experiences (White et al., 2016). For example, if an individual has learnt that one kind of emotion is more likely than others, they may have a biased expectation, and thus a biased start point, for responding "yes" to that emotion.

Table 1.

Typical parameter values of the drift-diffusion model as returned by fast-dm.(Voss, Voss, & Lerche, 2015)

Parameter	Typical range	Description
Boundary separation (*a*)	$0.5 < a < 2$	Larger values equate to a more conservative decision style
Relative Starting Point (z_r)	$0.3 < zr < 0.7$	Deviations from 0.5 indicate that different amounts of evidence are required to make a decision for the alternative responses.
Drift rate (*v*)	$-5 < v < 5$	The rate of evidence accumulation. Values further from zero indicate stronger evidence. Positive values indicate evidence for the upper threshold response, negative values indicate evidence for the lower threshold response.
Non-decision time (*t0*)	$0.1 < t0 < 0.5$	Average duration of all non-decisional processes (encoding and response execution).

The closer the starting point is to either boundary, the shorter the processing time will be for the corresponding response, and thus indicative of a decisional bias. An unbiased starting point will be returned as 0.5, indicating an equal distance from either decision boundary.

Drift rate (*v*). Drift rate is the mean rate of evidence accumulation toward one choice or the other, that is, the rate of approach to the decision threshold. A higher drift rate means faster accumulation of evidence. Drift rate is generally a function of stimulus or evidence quality. The easier an individual finds a stimulus to identify, the higher the drift rate. A higher drift rate for one emotion relative to others indicates a perceptual bias for that emotion.

Non-decision time (*t0*). The final parameter of relevance is non-decision time, *t0*, the time needed for sensory encoding prior to the evidence accumulation process and for response execution once a decision boundary has been crossed. This parameter is the component of response time that is not related to decisional processes.

Mapping Parameters onto Current Theory

Interpretation bias may be the result of numerous cognitive processes: perceptual bias, decisional bias, cognitive control, and/or non-decisional components of decision making. We hypothesised that perceptual and decisional biases are the most likely cognitive processes to underlie interpretation bias. To assess this, we examined the processes responsible for interpreting ambiguous stimuli in a non-clinical sample. We compared the fast-dm parameters returned from analysing two sets of data from the yes/no task. Data came from two groups of participants. One group was exposed to stimuli consisting of only clearly identifiable facial expressions, and the other group was exposed to clearly identifiable facial expressions intermixed with mild/ambiguous emotional expressions. Our goal was to identify which parameters are responsible for interpretation of affective stimuli of an ambiguous nature. Additionally, because cognitive control is related to inhibition difficulties, i.e., impulsiveness and difficulty regulating strong emotion, we explored the relationship between diffusion model parameters with scores on the difficulties with emotion regulation scale. Both the total score as well as the more specific impulsiveness sub scale was used for this purpose. From this, we aimed to explore the suitability of the yes/no task as a cognitive measure of emotion regulation difficulties.

We expected that increased cognitive control would be represented by a larger boundary separation, which we operationalised as a potential indicator of emotion regulation capacity. Given the non-clinical sample, we do not expect this parameter to differ between groups. However, as wider boundary separation indicates more conservative responding we expect decreased boundary separation to be associated with inhibited cognitive control, and thus to correlate with emotion regulation difficulties scores.

Due to the increased proportion of yes responses relative to no responses in the yes-no task, we expect to see a start point bias toward yes across all participants. This start point bias is expected to be moderated by the inclusion of ambiguous stimuli as the ambiguity of the expressions may drive down the subjective probability of a yes response in the emotion conditions, and conversely it may drive up the subjective probability of a yes response in the neutral conditions. As a result, differences in start point across conditions are indicative of a decisional bias for that emotion.

Drift rate represents the rate at which evidence accumulates for whether the stimulus belongs to the target group. As the evidence derived from the ambiguous stimuli is poorer, we expect that drift rates would be higher for the group that is presented with only stronger intensity emotional stimuli than the group that is also presented with ambiguous stimuli. A higher drift rate for one emotional condition relative to the other emotional conditions indicates a perceptual bias.

Given the non-clinical nature of the sample, there was no theoretical reason to expect differences between groups in non-decision time. However, non-decision time may correlate with emotion regulation difficulties if the emotional stimuli evoke a hyper-vigilant response in individuals with poor cognitive control/difficulty regulating emotion. Additionally, as negative interpretation bias has been linked to emotion regulation difficulties, we also explored the relationship between drift rate and starting point with scores on the difficulties with emotion regulation scale. As we expect interpretation bias to be captured by starting point and drift rate, and interpretation bias has been associated with emotion regulation difficulties, we expected that

starting point and drift rate for angry and sad emotional facial expressions would correlate with emotion regulation scores.

Method

This study received ethics approval from the Flinders University Social and Behavioural Research Ethics Committee (Project number 6259: Investigating cognitive mechanisms underpinning emotional cognitive biases in individuals with varying levels of emotional disorders).

Participant Characteristics and Exclusion Criteria

One hundred undergraduate women were recruited through the university's research participation program. Fifty participants received course credit for a first-year psychology topic, and fifty participants received $15 remuneration. Seven participants were excluded from the final data set for failing attention checks in the self-report measures. Thus, data from 93 participants, with a mean age of 22 years (min 17, max 47, median 20), were used in the final analysis.

Design

A 4 (emotion: happy, sad, angry, neutral) x 2 (response: yes, no) x 2 (stimuli: strong, mixed) mixed design was used, with stimulus type as the between subjects' factor. Dependent measures were fast-dm 30 parameters *a*, *sz*, *v*, and *t0*. The fast-dm 30 parameters for each emotional condition, *happy, sad, angry, neutral*, were used to explore their relationship with scores on the clinical measure of emotion regulation difficulties (total score and impulsiveness subscale).

Materials

In line with other research in this area (Arndt & Fujiwara, 2012; Niles et al., 2013; Paulewicz et al., 2012; Nim Tottenham et al., 2011), we selected facial stimuli from the NimStim database (Tottenham et al., 2009). This database consists of 646 facial expression stimuli across 43 faces of different gender and race. The database has expressions consisting of fearful, happy, sad, angry, surprised, calm, neutral, and disgusted, each with validity ratings (Tottenham et al., 2009). There are two sets of each expression. One set consists of the expressions with models' mouths open, and the other set consists of expressions with mouths closed. In this study, faces with angry,

happy, sad, and neutral expressions were selected from the set with mouths closed. The mouths-closed stimuli were chosen to ease the morphing process that was used to create a mild subset of happy, sad and angry expressions. Three faces were excluded as the models did not present with all four (angry, happy, sad, neutral) expressions. The remaining stimuli consisted of digital photographs of 40 individuals (18 females and 22 males), depicting each of the four expressions (happy, sad, angry, neutral). All pictures were presented in a black rectangular frame and cropped to just around the face to remove distinguishing external features, allowing the focus to be on the expression. All images were of equal size (332x458 pixels), resolution (120 pixels/inch), and RGB colour depth (8bpc). A subset of mild emotions was created by morphing faces of each valence with the corresponding neutral faces using the Fantamorph (Abrosoft, 2013) software program. These mild faces were morphed to a strength of 50%, exactly halfway between the original stronger intensity emotion and the corresponding neutral face.

Procedure

Data for the yes/no task were collected at the same time as data for another study investigating an alternative task, the dot probe task (Manuel, Kemps & McCarley, under review). Both tasks were completed on a Dell desktop computer with a standard QWERTY keyboard in one testing session that lasted approximately 45 minutes. The cognitive tasks (dot probe and yes/no task; counterbalanced) were presented first to avoid any mood induction that might result from the question content of the self-report measures. The yes/no task was run using Neurobehavioral Systems Inc. software. The clinical measures were collected in Qualtrics, an online survey program. The clinical measures were presented in the same order for each participant with the Difficulties with Emotion Regulation Scale (DERS) presented first, followed by the Depression, Anxiety and Stress Scale (DASS); only the DERS was used in this study.

Interpretation Bias/cognitive control: Yes/no task. (Tottenham et al., 2011). The yes/no task consisted of four blocks of 80 trials. For each trial, a single face appeared centred horizontally and vertically on a black background, and remained until the participant made a response indicating

that the expression did or did not match a target emotion. Participants were instructed to respond as quickly and accurately as possible. Response keys were the up arrow for "yes" and the down arrow for "no", labelled accordingly. A 500ms inter-stimulus interval followed the participant's response, after which the next trial began automatically.

At the beginning of each block, participants were asked, "Are the following faces neutral/happy/sad/angry?" The stimulus face matched the target valence on 2/3 (56 trials) of the trials within each block, to encourage a "yes" bias and the subsequent load on cognitive control when needing to inhibit the "yes" response on presentation of a distractor stimulus. On the remaining trials (24 trials), the stimulus faces expressed one of the three distractor emotions for that block. For example, if happy was the target emotion, distractor emotions were angry, sad and neutral. Target and distractor stimuli were generated randomly using a randomization algorithm. The order of the blocks was counterbalanced between participants.

The task had two stimulus sets. One group were shown emotional expressions of the original strength ("Strong" group; 48 participants), while the other group were shown emotional expressions that consisted of an equal mix of the original stronger intensity expressions and the morphed, milder, more ambiguous expressions ("Mixed" group; 45 participants).

Difficulties with emotion regulation scale (DERS; (Gratz & Roemer, 2004a). This 36-item self-report questionnaire consists of six dimensions of emotion regulation: (a) non-acceptance of emotional reactions; (b) difficulty engaging in goal-directed behaviour; (c) difficulty controlling impulses; (d) lack of emotional awareness; (e) limited access to emotion-regulation strategies; and (f) lack of emotional clarity. The DERS begins each item with the phrase, "When I'm upset…". Participants rate each item on a five-point Likert scale from "almost never" (1) to "almost always" (5). Eleven items are reverse scored. Both the total score and the difficulty controlling impulses subscale were used in this study (hereafter referred to as the impulsiveness subscale). The total score can range from 36 to 180, and the impulsiveness subscale score ranges from 6-30, with a higher score representing increased difficulty with emotion regulation. In non-clinical samples of

adults, the DERS has an average total score of 75-80 (Gratz & Roemer, 2004). There are no published norms for the subscales. The DERS has high internal consistency (α=.93), good test-retest reliability (ρ_I=.88), and satisfactory construct and predictive validity (Gratz & Roemer, 2004). Within this sample, Cronbach's alpha was .95.

Analysis

Data Cleaning and preparation.

Yes/No task. RTs less than 150ms and more than 3 standard deviations from each participant's mean RT for each response by condition and strength were removed, accounting for .02% of total responses. Fast-dm 30 was run in line with recommendations from Voss, Voss, and Lerche's (2015) tutorial paper. The fast-dm 30 control file was set to allow non-decision time, start point, and boundary separation to vary across participants. Drift rates for "yes" and "no" were mapped separately, with trials for the strong and ambiguous "yes" facial expressions for the mixed stimuli group also mapped separately. These values were combined to provide an average drift rate. Additional parameters examining variability in starting point, drift rate, and non-decision time can also be introduced into model analysis. While more parsimonious models can be preferable to more complex models (Lerche & Voss, 2016), it has been recommended to include non-decision time variability in the model in order to achieve stable parameter estimates (Voss et al., 2015; Lerche and Voss, 2016). Therefore, inter-trial variability of non-decision time (s_{t0}) was also allowed to vary across participants.

Model fit was assessed by the Kolmogorov–Smirnow (KS) statistic provided by the fast-dm 30 software. The KS statistic assesses whether the distribution of observed data differs significantly from a predicted distribution. The KS fit statistic returned by fast-dm 30 did not reveal any significant deviations between empirical and estimated RT distributions for any of the participants. As Voss et al. (2015) caution against relying solely on the fit statistic returned by fast-dm 30 and recommend plotting the data graphically. Comparison of empirical and predicted response time distributions for each expression (happy, angry, sad, neutral) were plotted, along with quantile

probability plots comparing empirical and predicted response times for the 25th, 50th, and 75th percentiles of each expression (happy, sad, angry). These plots are displayed in Appendix 2.

DERS. DERS responses were checked for skewness and kurtosis, and were within acceptable limits (+/- 2). There were no outliers.

Fast-dm 30 analysis. A mixed ANOVA was conducted to analyse differences in parameters between the two stimulus subsets. Bonferroni correction was applied, and where sphericity tests indicated it applicable, the Greenhouse Geisser correction was applied.

Relationship between fast-dm parameters and DERS. Correlations were used to explore whether fast-dm 30 parameters were associated with the clinical outcome measures of the DERS total score and impulsiveness subscale.

Results

Response Times

Response times for yes and no responses are displayed in Figure 2. A 4 (emotion: happy, sad, angry, neutral) x 2 (response: yes, no) x 2 (stimuli: strong, mixed) mixed ANOVA was used to examine the effect of stimulus strength on response times for each emotional expression and response type. There were main effects of emotion $F(2.69, 91) = 45.89$, $p < .001$, $\eta^2_{partial} = .34$, and response $F(1, 91) = 51.48$, $p < .001$, $\eta^2_{partial} = .36$, but not of stimuli $F(3, 91) = .24$, $p = .62$, $\eta^2_{partial} = .003$. Pairwise comparisons revealed that mean response times to stimuli in the happy condition were the fastest, with response times 73ms faster than responses in the neutral condition ($p = .001$), 169ms faster than the angry condition ($p < .001$), and 237ms faster than the sad condition ($p < .001$). Responses in the neutral condition were second fastest, on average 96ms faster than the angry condition ($p < .001$) and164ms faster than the sad condition ($p < .001$). Responses in the angry condition were an average of 68ms faster than the sad condition ($p = .03$), with responses in the sad condition the slowest. Pairwise comparisons for response type revealed that “yes” responses were an average of 94ms slower than “no” responses ($p < .001$). Additionally, there was a significant 3 way interaction $F(2.74, 249.66) = 4.54$, $p = .005$, $\eta^2_{partial} = .05$. Post hoc analyses using a series of

independent samples t-tests showed that the only group difference was for "yes" responses in the neutral condition. The group that had been presented with a mix of stronger intensity and

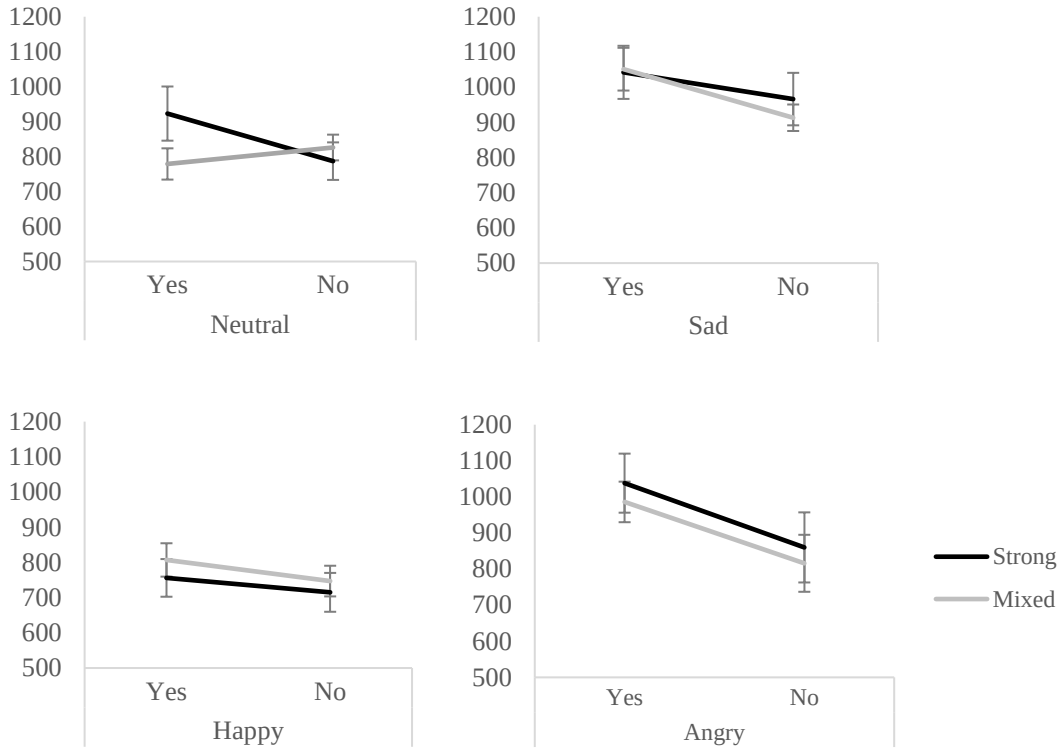


Figure 2. Mean response times (ms) for each stimulus type (strong, mixed) and emotion (neutral, sad, happy, angry); error bars are within subjects' standard errors.

Strong *n*=48, Mixed *n*=45

ambiguous expressions among the distractor stimuli were significantly faster (*M* = 779, *SD* = 174) to respond "yes" than the group that had been presented with only stronger intensity distractor stimuli (*M* = 923, *SD* = 347), *t*(70.16) = 2.57, *p* = .012, *d = 0.52.* There were no significant differences between groups for the other emotion-response combinations (all *p* values >.05). Thus, the inclusion of ambiguous information did not appear to affect response times in this sample.

Diffusion Model Parameters

Boundary separation (a). A 4 (emotion: happy, sad, angry, neutral*)* x 2 (stimuli: strong, mixed) mixed ANOVA was used to examine differences between boundary separation values (*a*; figure 3) across emotions and stimulus types. There was no effect of emotion, *F*(3, 273) = 1.99, *p* =

.12, $\eta^2_{partial}$ = .02, indicating no significant difference for *a* across emotions. However, there was a main effect of stimulus type, $F(1,91) = 3.94$, $p = .05$, $\eta^2_{partial} = .04$, indicating that boundary separation was smaller, and thus decisions were less conservative with the inclusion of ambiguous expressions ($M = 1.55$, $SD = .34$) relative to clearer, more intense expressions ($M = 1.68$, $SD = .45$). There was a significant interaction between emotion and stimulus type, $F(3,273) = 5.33$, $p = .001$, $\eta^2_{partial} = .06$. Independent samples t-tests identified that stimulus type only influenced the identification of neutral faces, such that boundary separation was smaller, and thus less conservative, for the group presented with both ambiguous and strong expressions ($M = 1.42$, $SD = .25$), than the group presented with only strong expressions ($M = 1.70$, $SD = .49$, $t(71.58) = 3.56$, $p = .001$, $d = 0.72$). There was no difference between groups for happy (mixed $M = 1.64$, $SD = .41$, strong $M = 1.60$, $SD = .41$, $t(91) = 0.52$, $p = .61$, $d = 0.10$), sad (mixed $M = 1.58$, $SD = .35$, strong $M = 1.73$, $SD = .45$, $t(91) = 1.74$, $p = .09$, $d = 0.37$), or angry (mixed $M = 1.54$, $SD = .35$, strong $M = 1.70$, $SD = .47$, $t(86.71) = 1.79$, $p = .08$, $d = 0.39$) emotional expressions. These results indicate that the inclusion of ambiguous stimuli resulted in less conservative decision-making in the neutral condition. This was contrary to our expectation that stimuli strength would not impact boundary separation.

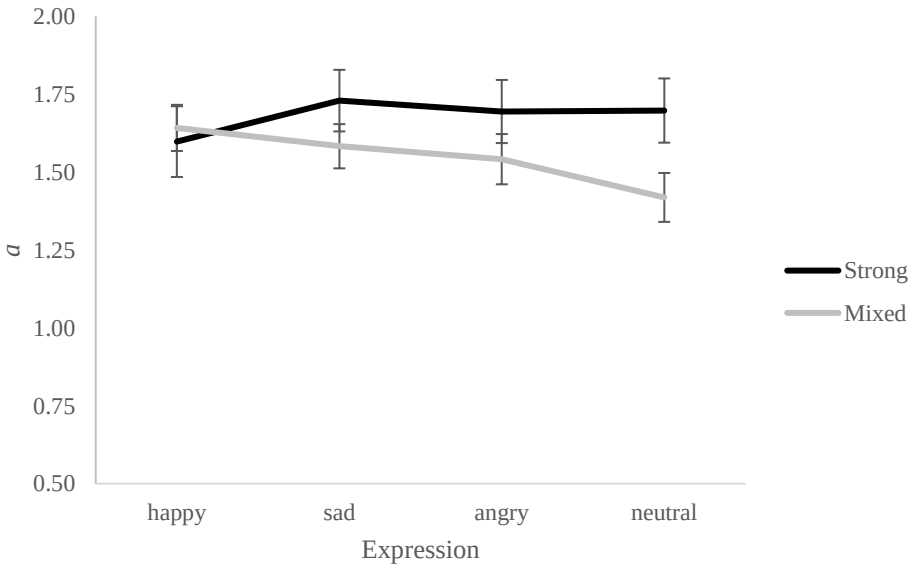


Figure 3. Boundary separation for each emotion by stimulus type; error bars are within subjects' standard errors.

Relative starting point (zr). A 4 (emotion: happy, sad, angry, neutral*)* x 2 (stimuli: strong, mixed) mixed ANOVA was used to examine differences between relative starting point values (*zr*; figure 4) for each expression and stimulus type. A main effect of emotion, $F(3,273) = 20.27$, $p < .001$, $\eta^2_{partial} = .18$, indicated a significant difference for *zr* across emotions. Pairwise comparisons revealed that the mean relative starting point value for the happy condition ($M = .65$, $SD = .10$) was higher than for the neutral ($M = .53$, $SD = .13$, $p < .001$) and sad ($M = .60$, $SD = .11$. $p = .40$) conditions, indicating a "yes" bias for happy relative to neutral and sad facial expressions. Additionally, the starting points for the sad and angry ($M = .63$, $SD = .13$) conditions were also significantly higher than for the neutral condition (both p values < .001). There were no significant differences between the starting points of happy and angry ($p = 1.00$), or sad and angry ($p = .37$) emotional expressions. There was a main effect of stimulus type, (mixed $M = .56$, $SD = .12$, strong $M = .62$, $SD = .11$), $F(1,91) = 5.89$, $p = .02$, $\eta^2_{partial} = .06$, indicating that the starting point for the strong expression group was higher than the mixed expression group. The interaction between emotion and stimulus type approached significance, $F(3,273) = 2.45$, $p = .06$, $\eta^2_{partial} = .03$. Independent samples t-tests identified a significant difference between groups for the sad condition, $t(91) = 2.15$, $p = .03$, approaching significance for the angry, $t(91) = 1.90$, $p = .06$, and neutral, $t(91) = 1.98$, $p = .051$, conditions, with no difference between groups in the happy condition, $t(91) = 1.10$, $p = .27$. Thus, starting point differed across expressions and between stimulus type, and indicated a trend toward those differences being more prominent for the sad, angry, and neutral conditions.

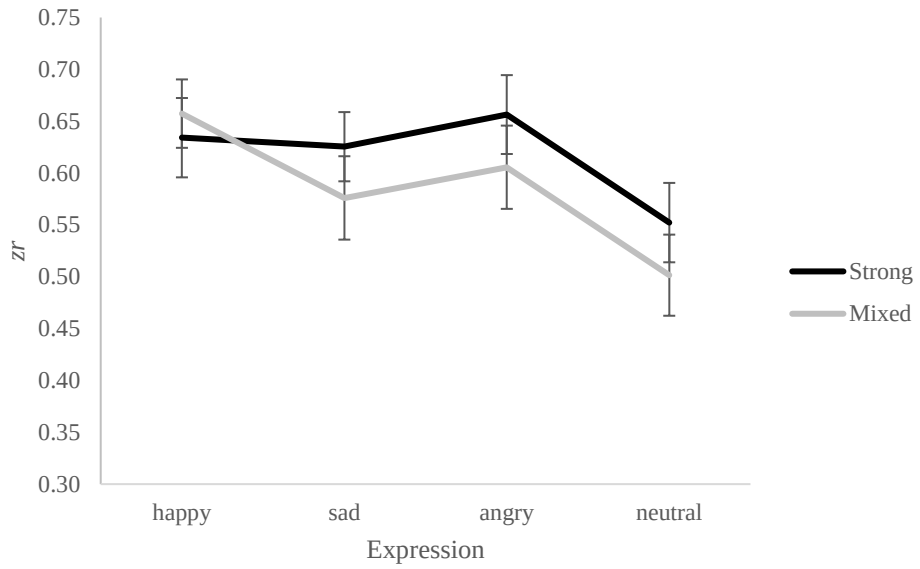


Figure 4. Relative starting point for each emotion by stimulus strength; error bars are within subjects' standard errors.

Drift rate (v). A 4 (emotion: happy, sad, angry, neutral*)* x 2 (stimuli: strong, mixed) mixed ANOVA was used to examine differences between drift rate values (*v*, figure 5) for each expression and stimulus type. A main effect of emotion, $F(3, 273) = 21.23$, $p < .001$, $\eta^2_{partial} = .19$, indicated a significant difference for *v* across emotions. Pairwise comparisons revealed that happy ($M = -0.49$, $SD = 0.96$) was easier to identify than angry ($M = -0.84$, $SD = 0.78$, $p = .03$), neutral ($M = 0.02$, $SD = 0.82$) was easier to identify than happy ($p < .001$), sad ($M = -0.31$, $SD = 0.79$) was easier to identify than angry ($p < .001$), but more difficult than neutral ($p = .01$), and neutral was easier to identify than angry ($p < .001$). Only the drift rates for happy and sad did not differ ($p = .91$). One sample t-tests revealed that the drift rates for happy, $t(92) = 4.88$, $p < .001$, sad, $t(92) = 3.83$, $p < .001$, and angry, $t(92) = 10.30$, $p < .001$, were all significantly different from zero, indicating that drift rate was stronger for distractor stimuli than the target stimuli. The drift rate for the neutral condition did not significantly differ from zero, $t(92) = 0.18$, $p = .86$, indicating that there was no difference in drift rate for emotional distractor stimuli or neutral stimuli.

There was no main effect of stimulus type, $F(1,91) = 0.02$, $p = .90$, $\eta^2_{partial} < .001$, but there was a significant interaction between emotion and stimulus type, $F(3, 273) = 8.80$, $p < .001$, $\eta^2_{partial} = .09$. Independent samples t-tests identified that drift rate differed between groups for the neutral condition, $t(91) = 4.93$, $p < .001$, but not for happy, $t(91) = 1.84$, $p = .07$, sad, $t(91) = 1.16$, $p = .25$, or angry, $t(80.95) = 0.50$, $p = .62$, signifying that the inclusion of ambiguous stimuli only impacted the neutral condition. To further examine the impact of the inclusion of ambiguous expressions on participants' ability to identify the more intense expressions, a 3 (emotion: happy, sad, angry) x 2 (stimuli: strong, mixed) mixed ANOVA was used to compare the drift rate for only the "strong" expressions in each group (figure 6). There was no main effect of group, $F(1.73, 157.38) = 1.08$, $p = .33$, $\eta^2_{partial} = .01$, indicating that drift rates for the strong stimulus "yes" trials did not differ between groups. However, there was a significant interaction, $F(1, 91) = 12.16$, $p = .001$, $\eta^2_{partial} = .12$. Independent samples t-tests identified that in the happy condition drift rates were faster for the intense emotions in the mixed stimuli group ($M = 3.27$, $SD = 1.87$) than in the strong stimuli group ($M = 2.43$, $SD = 1.21$), $t(91) = 2.58$, $p = .01$, $d = 0.53$. There was no difference between groups for drift rates in the sad (mixed $M = 1.74$, $SD = 1.17$, strong $M = 1.37$, $SD = 0.81$, $t(77.71) = 1.74$, $p = .09$, $d = 0.37$) or angry (mixed $M = 1.42$, $SD = 1.18$, strong $M = 1.03$, $SD = 0.97$, $t(91) = 1.78$, $p = .08$, $d = 0.36$) conditions. This suggests that the inclusion of ambiguous stimuli seems to have a rebound effect – making the "strong" stimuli easier to identify.

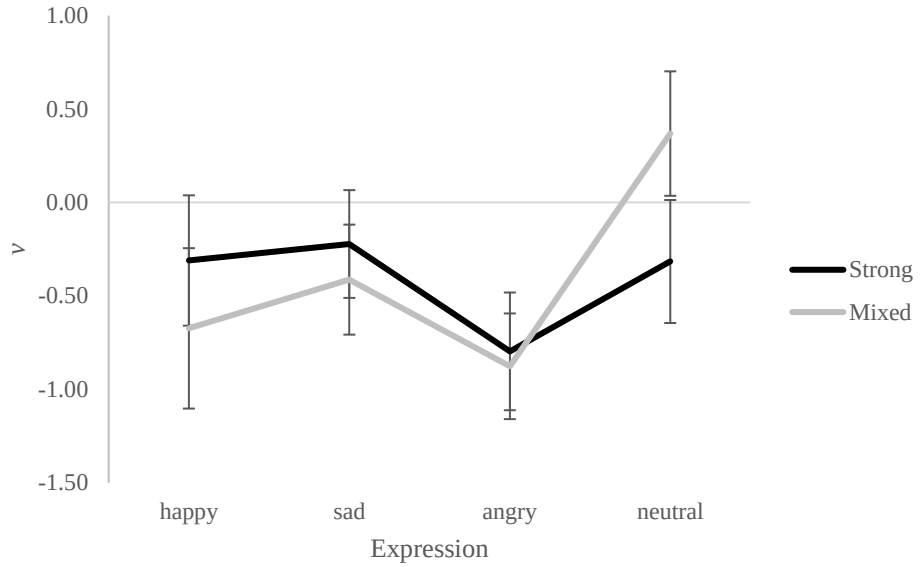


Figure 5. Drift rates for each emotion; error bars are within subjects' standard errors.

Figure 6. Drift rates for the strong expression stimuli within each group; error bars are within subjects' standard errors.

Non-decision time (t0). A 4 (emotion: happy, sad, angry, neutral*)* x 2 (stimuli: strong, mixed) mixed ANOVA was used to examine differences between non-decision time values (*t0,* figure 7) for each expression and stimulus type. A main effect of emotion, $F(2.75, 250.13) = 12.30$, $p < .001$, $\eta^2_{partial} = .12$, indicated a significant difference for *t0* across emotions. Pairwise comparisons revealed that non-decision time for the happy condition ($M = .43$, $SD = .07$) was significantly faster than for the sad ($M = .48$, $SD = .10$, $p < .001$), angry ($M = .48$, $SD = .09$, $p < .001$), and neutral ($M = .49$, $SD = .08$, $p < .001$) conditions. However, non-decision time did not significantly differ between

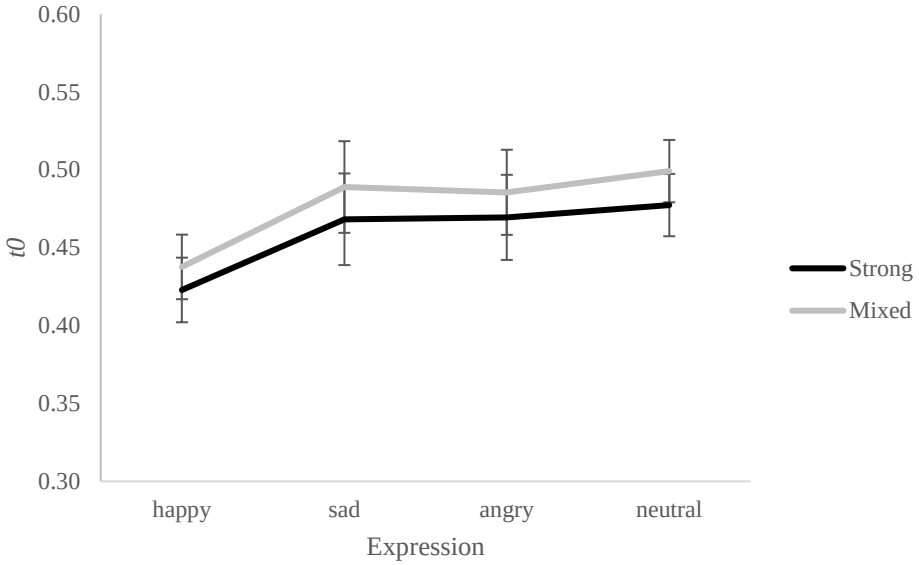


Figure 7. Non-decision time for each emotion; error bars are within subjects' standard errors.

neutral and angry, neutral and sad, or sad and angry expressions (all *p* values = 1.00). There was no main effect of stimulus type, $F(1, 91) = 2.03$, $p = .16$, $\eta^2_{partial} = .02$, nor was there a significant interaction between emotion and stimulus type, $F(2.75, 250.13) = 0.05$, $p = .98$, $\eta^2_{partial} = .001$. Therefore, non-decision time was faster when identifying happy expressions than it was when identifying sad, angry, or neutral expressions.

Clinical Measures. The mean total score of the DERS ($M = 92.60$, $SD = 23.52$) was slightly above the non-clinical range of 75-80 identified by Gratz and Roemer (2004), indicating that the sample included a range of individuals with clinically significant scores. The mean score of the impulsiveness subscale ($M = 14.10$, $SD = 5.53$) was around the mid-point of the scale, indicating that impulsiveness difficulties across the sample were reported to occur "about half the time". As can be seen in table 2, the total score on the DERS was negatively correlated with boundary separation, indicating that a more conservative response criterion in the face of angry stimuli is indicative of better emotion regulation skills; however, this result was no longer significant

Table 2

Correlations between fast-dm parameters and DERS total score and impulsiveness subscale

	Total score				Impulsiveness			
	Happy	Sad	Angry	Neutral	Happy	Sad	Angry	Neutral
z_r	-0.08	0.05	-0.01	-0.06	0.09	0.10	-0.02	-0.06
a	-0.17	-0.13	-0.23*	-0.15	-0.34**	-0.28**	-0.36**	-0.27**
v_{yes}	0.03	0.08	0.27**	0.09	0.10	0.03	0.17	0.08
v_{no}	0.08	0.10	0.11	0.12	0.00	-0.03	0.05	0.08
t0	-0.01	-0.12	-0.05	-0.09	-0.09	-0.13	-0.04	-0.08

Note: * = significant at the .05 level; ** = significant at the .001 level.

after Bonferroni correction and thus must be taken with caution. By contrast, the boundary separation and DERS impulsiveness subscale were negatively correlated across all emotions, with the relationship between impulsiveness and the angry condition remaining significant after Bonferroni correction. These results indicate that more conservative responding may be related to impulse control.

Discussion

Biased interpretation of affective information and difficulty regulating emotion have been linked to the onset and maintenance of clinical disorders such as anxiety and depression. However, popular measures of interpretation bias that involve ambiguous written scenarios are unable to identify the cognitive processes that are responsible for interpretation bias. Further, one of the most commonly used measures of emotion regulation difficulties, the difficulties with emotion regulation scale (Gratz & Roemer, 2004a), is a self-report measure. Self-report measures are open to response

bias and imprecise responding from individuals who lack of awareness of their own emotions, and thus can have compromised validity. As such, our aim was to investigate the decision-making processes that are responsible for interpretation of ambiguous stimuli in a non-clinical sample by analysing data from an affective yes-no task using the diffusion decision model (Ratcliff, 1978). We also explored the suitability of the decision-making parameters returned from the diffusion model analysis as an alternative cognitive measure of emotion regulation difficulties.

We hypothesised that perceptual bias and decisional bias would be the most likely candidates underlying interpretation bias. Therefore, we expected that the inclusion of ambiguous stimuli would affect start point and drift rate, but not boundary separation or non-decision time. Our expectations were supported in that both start point and drift rate were affected by stimulus type. The increased ratio of "yes" to "no" trials resulted in a start point (decisional) bias toward "yes" for the emotion conditions in the strong stimuli group, indicating the stimulus ratio produced the effect that was expected.

Further, we expected that the increased ambiguity of the emotional expressions might reduce the perceived ratio of "yes" to "no" responses in the emotional conditions, and increase the perceived ratio of "yes" to "no" responses in the neutral condition. There were indication of this trend in the sad and angry conditions, as the inclusion of ambiguous stimuli reduced the start point bias in the sad and angry conditions, but it did not affect the bias in the happy condition, and the reverse effect was not found for the neutral condition. The relatively higher start point in the happy condition for the mixed stimuli group indicates a decisional bias for positive emotion in that group. This indicates that ambiguous stimuli were interpreted more positively in this condition. As this study used a non-clinical sample, this finding is consistent with the content specificity hypothesis (Beck, 1976) which states that non-clinical individuals tend to interpret ambiguous information in a positive light, whereas individuals with depression are more likely to interpret information in a dysphoric manner, and those with anxiety are more likely to interpret information in a threatening manner.

We expected that the presence of a perceptual bias, as identified by the drift rate, would be more noticeable when using the ambiguous stimuli due to their equivocal categorisation. The inclusion of ambiguous expressions resulted in lower drift rates for the mixed group than the stronger intensity group overall. Additionally, differences existed across emotion such that in the neutral condition, participants accumulated evidence for neutral expressions more quickly than they did for emotional expressions. Conversely, in the angry condition, participants accumulated evidence more slowly for angry expressions than the mix of happy, sad, and neutral distractor expressions. However, this pattern was not dependent on stimulus strength. This finding indicates that lower drift rates resulting from the addition of ambiguous emotive stimuli in a non-clinical sample offers the potential to reduce ceiling effects that can result from stronger intensity emotive stimuli. Therefore, in a clinical sample, interpretation bias may result in higher drift rates, enabling us to identify perceptual bias.

We did not expect boundary separation or non-decision time to differ between groups as a function of stimulus strength. This was not the case for boundary separation but was for non-decision time. Contrary to our expectations, boundary separation differed between groups in the neutral condition. There was no effect of stimulus strength in the other conditions, however, inclusion of ambiguous stimuli in the neutral condition reduced decision caution. Non-decision time did not differ between groups, however non-decision times for happy expressions were faster than non-decision times for the other expressions, all of which did not differ from each other. This was unexpected as we did not expect any difference in non-decision time across conditions. As this was a non-clinical sample, this faster non-decisional time for happy expressions, is consistent with the content specificity hypothesis. This suggests that preference for mood congruent stimuli may also be a function of non-decision processes.

As hypothesised, boundary separation was related to emotion regulation; however, this relationship was not strong. Increased impulsivity was related to less conservative responding across all emotions, and total difficulty with emotion regulation scores were related to smaller

boundary separation for the angry condition. While these relationships were found to exist, they were relatively weak, and many became non-significant after Bonferroni correction. There was no link between emotion regulation difficulties and starting point (decisional bias), suggesting that although a decisional bias toward happy emotional expressions was found, it was not a protective factor for emotion regulation difficulties. Further, drift rate parameters did not identify a link between perceptual bias and emotion regulation difficulties, and there was no link between non-decision time and emotion regulation difficulties. Thus, although there are benefits to a cognitive measure that can capture emotion regulation difficulties irrespective of a social desirability bias or limited emotional self-awareness, the yes-no task did not identify a link with self-reported emotion regulation difficulties in a non-clinical sample.

Taken altogether, drift diffusion analysis has identified the cognitive processes that underlie interpretation of ambiguous information in a way that RT analysis cannot. The inclusion of ambiguous stimuli identified a difference in both perceptual bias and decisional bias between the identification of stronger intensity and ambiguous expressions. This indicates that in a non-clinical sample, it appears that perceptual bias *and* decisional biases are affected by interpreting ambiguous expressions. Additionally, decision boundary and non-decision time are both affected by interpretation of different emotions irrespective of stimulus strength. Future research using clinical samples will need to confirm whether interpretation bias in individuals with anxiety and depression is based on perceptual bias, decisional bias, or both, and whether decision boundary and non-decision time play a role in clinical group differences.

As negative expressions are posited to be easier to discriminate for those with a negative bias, and ambiguous emotion is interpreted in a mood-congruent manner, this perceptual advantage may manifest in a perceptual bias. In this case, a threat bias (anxiety-related) may result in a higher drift rate for angry facial expressions than for sad or happy expressions, whereas and a dysphoric bias (depression-related) may result in a higher drift rate for sad facial expressions than for angry or happy expressions. Additionally, as individuals with anxiety and depression commonly have in their

history pertinent negative life events, they may possess an increased expectation that disorder-congruent affect is more prevalent, resulting in a decisional bias. In this case, the bias would be evident in a higher start point reflecting the individual's expectation of negative information. Diffusion model analysis would enable us to identify whether both perceptual and decisional biases underlie anxiety and depression, whether one is more prevalent for one disorder than another, or whether there are individual differences amongst individuals experiencing the disorders.

In conclusion, the current study has demonstrated that diffusion decision model analysis can differentiate between the cognitive processes responsible for interpretation of ambiguous stimuli, and therefore has the potential to identify the underlying processes of interpretation biases involved in the onset and maintenance of clinical disorders. Such knowledge will benefit researchers in offering a better understanding of the underlying cognitive processes, and facilitate the development of more targeted methods of bias modification. It would also aid clinicians to more effectively formulate their treatment plans and provide more tailored interventions.

References

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders-DSM-5. *Washington, DC: Author*.

Amir, N., Beard, C., Burns, M., & Bomyea, J. (2009). Attention modification program in individuals with generalized anxiety disorder. *Journal of Abnormal Psychology*, *118*(1), 28.

Arndt, J. E., & Fujiwara, E. (2012). Attentional bias towards angry faces in trait-reappraisal. *Personality and Individual Differences*, *52*(1), 61–66. https://doi.org/10.1016/j.paid.2011.08.030

Bar-Haim, Y. (2010). Research review: attention bias modification (ABM): a novel treatment for anxiety disorders. *Journal of Child Psychology and Psychiatry*, *51*(8), 859–870.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. International Universities Press.

Broadbent, D., & Broadbent, M. (1988). Anxiety and attentional bias: State and trait. *Cognition & Emotion*, *2*(3), 165–183.

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: results from a pilot study. *J Behav Ther Exp Psychiatry*, *42*, 258–64. https://doi.org/10.1016/j.jbtep.2010.12.006

Browning, M., Holmes, E. A., & Harmer, C. J. (2010). The modification of attentional bias to emotional information: A review of the techniques, mechanisms, and relevance to emotional disorders. *Cogn Affect Behav Neurosci*, *10*, 8–20. https://doi.org/10.3758/CABN.10.1.8

Burcusa, S. L., & Iacono, W. G. (2007). Risk for Recurrence in Depression. *Clinical Psychology Review*, *27*(8), 959–985. https://doi.org/10.1016/j.cpr.2007.02.005

Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016). Scaling-up treatment of depression and anxiety: a global return on investment analysis. *The Lancet. Psychiatry*, *3*(5), 415–424. https://doi.org/10.1016/S2215-0366(16)30024-4

Cisler, J. M., & Olatunji, B. O. (2012). Emotion Regulation and Anxiety Disorders. *Current Psychiatry Reports*, *14*(3), 182–187. https://doi.org/10.1007/s11920-012-0262-2

Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & Emotion*, *6*(6), 409–434.

Eysenck, M. W., 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*(2), 144.

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General*, *136*(3), 389.

Goodwin, L. D., & Leech, N. L. (2006). Understanding Correlation: Factors That Affect the Size of r. *The Journal of Experimental Education*, *74*(3), 249–266. https://doi.org/10.3200/JEXE.74.3.249-266

Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: current status and future directions. *Annual Review of Clinical Psychology*, *6*, 285.

Gratz, K. L., & Roemer, L. (2004a). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment*, *26*(1), 41–54.

Gratz, K. L., & Roemer, L. (2004b). Multidimensional Assessment of Emotion Regulation and Dysregulation: Development, Factor Structure, and Initial Validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment*, *26*(1), 41–54. https://doi.org/10.1023/B:JOBA.0000007455.08539.94

Gross, J. J. (2013). Emotion regulation: Taking stock and moving forward. *Emotion*, *13*, 359–365. https://doi.org/10.1037/a0032135

Gross, J. J., & Thompson, R. A. (2007). Emotion regulation: Conceptual foundations. Retrieved from http://psycnet.apa.org/psycinfo/2007-01392-001

Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychol Bull*, *137*, 940–58. https://doi.org/10.1037/a0024355

Heathcote, A., Suraev, A., Curley, S., Gong, Q., Love, J., & Michie, P. T. (2015). *Decision processes and the slowing of simple choices in schizophrenia.* American Psychological Association. Retrieved from http://psycnet.apa.org/journals/abn/124/4/961/

Joormann, J., & Quinn, M. E. (2014). Cognitive processes and emotion regulation in depression. *Depress Anxiety*, *31*(4), 308–15. https://doi.org/10.1002/da.22264

Joormann, Jutta, & Siemer, M. (2011). Affective processing and emotion regulation in dysphoria and depression: Cognitive biases and deficits in cognitive control. *Social and Personality Psychology Compass*, *5*(1), 13–28.

Lerche, V., & Voss, A. (2016). Model Complexity in Diffusion Modeling: Benefits of Making the Model More Parsimonious. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01324

Leyman, L., De Raedt, R., Schacht, R., & Koster, E. H. W. (2007). Attentional biases for angry faces in unipolar depression. *Psychological Medicine*, *37*(3), 393–402. https://doi.org/10.1017/S003329170600910X

Lovibond, S. H., & Lovibond, P. F. (1996). *Manual for the depression anxiety stress scales*. Sydney : Psychology Foundation of Australia. Retrieved from http://trove.nla.gov.au/version/38629420

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, 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. https://doi.org/10.1037//0021-843x.111.1.107

Mogg, K., & Bradley, B. P. (1999). Some methodological issues in assessing attentional biases for threatening faces in anxiety: A replication study using a modified version of the probe detection task. *Behaviour Research and Therapy*, *37*(6), 595–604.

Mogg, K., & Bradley, B. P. (2016). Anxiety and attention to threat: Cognitive mechanisms and treatment with attention bias modification. *Behaviour Research and Therapy*, *87*(Supplement C), 76–108. https://doi.org/10.1016/j.brat.2016.08.001

Niles, A. N., Mesri, B., Burklund, L. J., Lieberman, M. D., & Craske, M. G. (2013). Attentional bias and emotional reactivity as predictors and moderators of behavioral treatment for social phobia. *Behaviour Research and Therapy*, *51*(10), 669–679. https://doi.org/http://dx.doi.org/10.1016/j.brat.2013.06.005

Nosek, B. A., & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition*, *19*(6), 625–666. https://doi.org/10.1521/soco.19.6.625.20886

O'Leary, M. J., & Barber, P. J. (1993). Interference effects in the Stroop and Simon paradigms. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(4), 830.

Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*, *5*(5), 376–404. https://doi.org/10.1167/5.5.1

Paulewicz, B., Blaut, A., & Kłosowska, J. (2012). Cognitive effects of attentional training depend on attentional control. *Polish Psychological Bulletin*, *43*(4), 272–277. https://doi.org/10.2478/v10059-012-0030-5

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.

Renaud, J., Russell, J. J., & Myhr, G. (2014). Predicting Who Benefits Most From Cognitive-Behavioral Therapy for Anxiety and Depression. *Journal of Clinical Psychology*, *70*(10), 924–932. https://doi.org/10.1002/jclp.22099

Rodebaugh, T. L., Scullin, R. B., Langer, J. K., Dixon, D. J., Huppert, J. D., Bernstein, A., … Lenze, E. J. (2016). Unreliability as a threat to understanding psychopathology: The

cautionary tale of attentional bias. *Journal of Abnormal Psychology, 125*(6), 840–851. https://doi.org/10.1037/abn0000184

Schoenmakers, T., Wiers, R. W., Jones, B. T., Bruce, G., & Jansen, A. T. M. (2007). Attentional re-training decreases attentional bias in heavy drinkers without generalization. *Addiction (Abingdon, England), 102*(3), 399–405. https://doi.org/10.1111/j.1360-0443.2006.01718.x

Scholten, W. D., Batelaan, N. M., van Balkom, A. J., WJH. Penninx, B., Smit, J. H., & van Oppen, P. (2013). Recurrence of anxiety disorders and its predictors. *Journal of Affective Disorders, 147*(1), 180–185. https://doi.org/10.1016/j.jad.2012.10.031

Schulz, K. P., Fan, J., Magidina, O., Marks, D. J., Hahn, B., & Halperin, J. M. (2007). Does the emotional go/no-go task really measure behavioral inhibition?: Convergence with measures on a non-emotional analog. *Archives of Clinical Neuropsychology, 22*(2), 151–160. https://doi.org/10.1016/j.acn.2006.12.001

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. *J Abnorm Psychol, 118*, 65–75. https://doi.org/10.1037/a0014377

Shechner, T., Britton, J. C., Pérez-Edgar, K., Bar-Haim, Y., Ernst, M., Fox, N. A., … Pine, D. S. (2012). Attention biases, anxiety, and development: toward or away from threats or rewards? *Depression and Anxiety, 29*(4), 282–294. https://doi.org/10.1002/da.20914

Taylor, C. T., Bomyea, J., & Amir, N. (2011). Malleability of attentional bias for positive emotional information and anxiety vulnerability. *Emotion, 11*(1), 127.

Tottenham, N., Hare, T. A., & Casey, B. J. (2011). Behavioral assessment of emotion discrimination, emotion regulation, and cognitive control in childhood, adolescence, and adulthood. *Front Psychol, 2*, 39. https://doi.org/10.3389/fpsyg.2011.00039

Tottenham, Nim, Hare, T. A., & Casey, B. (2011). Behavioral assessment of emotion discrimination, emotion regulation, and cognitive control in childhood, adolescence, and adulthood. *Frontiers in Psychology, 2*.

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods, 39*(4), 767–775. https://doi.org/10.3758/BF03192967

Voss, A., Voss, J., & Lerche, V. (2015a). Assessing cognitive processes with diffusion model analyses: a tutorial based on fast-dm-30. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00336

Voss, A., Voss, J., & Lerche, V. (2015b). Assessing cognitive processes with diffusion model analyses: a tutorial based on fast-dm-30. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00336

Waters, A. M., Pittaway, M., Mogg, K., Bradley, B. P., & Pine, D. S. (2013). Attention training towards positive stimuli in clinically anxious children. *Developmental Cognitive Neuroscience, 4*, 77–84. https://doi.org/10.1016/j.dcn.2012.09.004

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Using diffusion models to understand clinical disorders. *J Math Psychol, 54*, 39–52. https://doi.org/10.1016/j.jmp.2010.01.004

White, Corey N., Skokin, K., Carlos, B., & Weaver, A. (2016). Using decision models to decompose anxiety-related bias in threat classification. *Emotion, 16*(2), 196.

Zvielli, A., Bernstein, A., & Koster, E. H. (2015). Temporal dynamics of attentional bias. *Clinical Psychological Science, 3*(5), 772–788.

Chapter 5: Study 3

Evidence accumulation analysis of attentional bias modification in anxiety and depression: A failure to find attentional bias

Amy Manuel, Eva Kemps, Jason S. McCarley

School of Psychology, Flinders University, Adelaide, Australia

Author Note

Jason S McCarley, School of Psychology, Flinders University, is now at the College of Liberal Arts, Oregon State University, USA

Correspondence concerning this article should be addressed to Amy Manuel, School of Psychology, Flinders University, GPO Box 2100, Adelaide, SA 5001, Australia. Electronic mail may be sent to Amy.Manuel@flinders.edu.au

Abstract

Efforts to modify cognitive biases in anxiety and depression to alleviate symptoms is a growing area of research. However, the cognitive processes responsible for successful bias modification are still not well understood. Therefore, the aim of this study was to use diffusion model analysis to identify which cognitive processes are affected by attentional bias modification training. An intervention study was carried out to assess the cognitive processes that underlie successful attentional bias modification for participants with a history of anxiety and depression, relative to those without. Participants (N = 252) performed an active or sham training paradigm, and their scores of attentional bias (using a dot probe task), and interpretation bias (using a yes/no task), were compared pre and post training. Results did not identify an attentional or interpretation bias at pre- training, and as such, there were no effects of training found. Therefore, the ability to identify the processes responsible for successful bias modification was impeded. Diffusion model analysis did identify differences in decisional processes for individuals with a history of anxiety and depression, relative to those without. The findings presented in this paper support the growing notion that attentional bias may not be a stable trait. To identify the processes responsible for successful attentional bias modification, recommendations are made to apply diffusion model analysis to data from an attentional bias modification dataset that has already established a successful change in attentional bias using mean RTs, and examine the effects of training that way.

Key words: Attentional Bias Modification (ABM), Diffusion decision model, Anxiety, Depression, Fast-dm 30

Evidence accumulation analysis of attentional bias modification in anxiety and depression: A failure to find attentional bias

Anxiety and depression have the highest prevalence rates of all mental health disorders, and have high rates of relapse (Burcusa & Iacono, 2007; Scholten et al., 2013). One of the difficulties in treating these disorders has been attributed to the existence of implicit cognitive biases for disorder-relevant information, which have long been posited to contribute to the onset and maintenance of anxiety and depression (Beck, 1976). Beck's content specificity hypothesis states that anxiety is associated with an implicit preferential attention for threat-related information (attentional bias), and interpretation of ambiguous information in a threatening manner (interpretation bias). Similarly, depression is associated with an attentional bias for dysphoric information, and a dysphoric interpretation bias (Beck, 1976). These biased cognitive processes are hypothesised to exacerbate disorder-relevant symptoms through increasing emotion regulation difficulties (Cisler & Olatunji, 2012; Joormann & Quinn, 2014).

Emotion regulation is the ability to monitor, evaluate, and modify the magnitude and duration of one's emotional reactions (Svaldi et al., 2012). Being able to regulate strong emotion is an important part of social functioning and good mental health (Gross, 2013). The inability to regulate strong emotional reactions can result in impulsive responding, which is detrimental to social connectedness. Reacting impulsively to emotion can alienate others and increase feelings of isolation (Gross, 2013). As a result, these experiences of increased negative emotion can exacerbate disorder-relevant symptoms, leading to the onset and maintenance of affective disorders such as anxiety and depression (Cisler & Olatunji, 2012; J. Joormann & Quinn, 2014).

Attentional Bias Modification

Because of the hypothesised role of cognitive biases in the development and maintenance of anxiety and depression, researchers have turned to exploring ways to modify these biases. The goal of bias modification is to alleviate the affective symptoms of the targeted psychological disorders. One form of modification is known as attentional bias modification (MacLeod et al., 2002). The

goal of ABM is to shift implicit attention away from disorder-congruent negative information. ABM is hypothesised to be able to train individuals to implicitly attend away from negative stimuli, and thus reduce their attention to disorder-congruent stimuli. This shift in attention is hypothesised to alleviate the negative affective experiences that are associated with attending to negative information. It is hypothesised that by disrupting an individual's attention to negative information, the subsequent process of interpretation is also modulated, and thus the strong emotions that commonly follow from negative attention and interpretation are prevented (Gross & Thompson, 2007).

Some research findings indicate that ABM has therapeutic effects on the symptoms of both anxiety and depression (e.g., Amir, Beard, Burns, & Bomyea, 2009; Brosan, Hoppitt, Shelfer, Sillence, & Mackintosh, 2011; See, MacLeod, & Bridle, 2009), and research has started to explore whether training individuals away from negative stimuli, or training them toward happy stimuli is more effective (e.g., Shechner et al., 2012; Taylor, Bomyea, & Amir, 2011; Waters, Pittaway, Mogg, Bradley, & Pine, 2013). This research is in its infancy, but may nevertheless offer interesting insight into training efficacy. Bar-Haim (2010) posited that while there is evidence for success of training away from negative information, there is evidence of deficits in processing of positive information, and therefore recommends the exploration of training toward positive stimuli to assess the therapeutic effects.

The ABM training paradigm is derived from the same task that is primarily used to measure attentional bias, the dot probe task. This computerised task instructs participants to focus on a fixation cross in the centre of a computer screen. The fixation interval is followed by the brief presentation of two stimuli, typically images or words, side-by-side or vertically aligned. One stimulus is emotional and the other is neutral in valence. In the task's simplest form, after the stimuli disappear, a dot appears in the location previously occupied by one of the stimuli, and the participant is required to execute a speeded response indicating in which of the two positions the probe appeared (Koster, Crombez, Verschuere, & De Houwer, 2004; MacLeod, Mathews, & Tata,

1986). In alternative versions of the task, different probes are used, e.g., two letters, or two arrows pointing in opposite directions, and participants are required to identify which probe appears (Mogg & Bradley, 1999). The participant is instructed to ignore the stimuli and focus only on the location/identity of the probe. Trials on which the probe replaces the emotional stimulus are called congruent trials, and trials on which the probe replaces the neutral stimulus are called incongruent trials. There are equal numbers of congruent and incongruent trials for all emotional categories across the task. If the participant's attention has been captured by a specific emotional stimulus, it is assumed to be reflected by faster RTs when the probe appears in the location previously occupied by that stimulus. An estimate of attentional bias for emotional stimuli is calculated by subtracting the mean RT for congruent trials from the mean RT for incongruent trials. Positive scores result from faster RTs to congruent trials, and are taken to indicate an attentional bias toward that emotion.

The training version of the dot probe task is similar, and was devised by MacLeod and colleagues (2002). Task instructions are the same, however, the range of stimuli is limited to the stimulus type being trained toward, e.g., neutral/happy, and the stimulus type being trained away from, e.g., negative/neutral. While in the assessment version of the dot probe task, probes are presented equally often following each stimulus type, in the training variant, the probe appears most often at the location of the stimulus type that is the focus of training (e.g., neutral/happy), and only infrequently at the location of the stimulus type that is the target of modification (e.g., negative/neutral). For example, if training away from negative stimuli, the probe will appear behind the neutral stimulus on most of the trials and only appears behind the negative stimulus on a minimal number of trials to inhibit awareness of training contingency. Similarly, if training toward happy stimuli, the probe will appear behind the happy stimuli on most of the trials, and behind the neutral stimulus on a minimal number of trials. This format is hypothesised to implicitly train attention toward the target emotion.

To ensure subsequent changes in bias scores can be attributed to this training paradigm, a control training condition (sham training) is implemented. The sham training consists of the same stimuli as the training paradigm, and the same number of trials, however the probe is presented behind each stimulus type with equal frequency. Following this protocol, shifts in attentional bias in the active training group that differ from the sham training group, are hypothesised to be a function of the training contingency. While numerous reviews in the past decade have praised ABM as a potentially effective treatment for anxiety and depression (e.g., Baert, Koster, & De Raedt, 2011; Hertel & Mathews, 2011; MacLeod, 2012; MacLeod & Mathews, 2012), how exactly successful bias modification works remains unclear. It is hypothesised that training attention toward or away from a given emotion works at an implicit level, however the current attentional bias score can only reveal that something has happened, not what has happened or when (Cisler et al., 2009).

Another, and more concerning limitation of the dot probe task is that the psychometric properties of the attentional bias score are inadequate (Schmukle, 2005; Cisler et al., 2009; Price et al., 2015). Specifically, the attentional bias score has low internal consistency (split half and Cronbach's alpha) and poor test-retest reliability (see Schmukle, 2005). Without adequate psychometric rigour, the data derived from the dot probe task can only be presented with a level of confidence on par with the properties of the task. In intervention studies specifically, reliability of a measure is a particularly important factor. The reliability of the measure used to assess outcomes is required to be much higher than for research identifying statistical association or differences between groups (Rodebaugh et al., 2016).

MacLeod and Clarke (2015) have argued that while the dot probe task is not a satisfactory measure of inter-individual differences, the dot probe task can reliably detect group differences (e.g., clinical versus non-clinical; Bar-Haim et al., 2007) as well as the success of ABM training conditions (e.g., active versus placebo training; Hakamata et al., 2010). Price and colleagues (2015), while acknowledging the instances where the dot probe task did differentiate between groups, noted the published exceptions where no group differences were found (e.g., Mohlman, Price, & Vietri,

2013; Price et al., 2013; Waters, Lipp, & Spence, 2004). They further highlighted the risk of a file drawer effect, which means that there are likely numerous studies that did not find differences and thus were not published. In addition to this, Rodebaugh et al. (2016) point out that most attentional bias modification studies do not present reliability coefficients of the task with the results, and those that do fall well short of the recommended level of at least .90, hovering around .45 (see Rodebaugh et al., 2016, for a review). Without adequate reliability of the bias measure, confidence in concluding that a third variable (i.e., the modification training) was the mechanism for change is compromised. Specifically, if the reliability of the attentional bias measure cannot be established, alternative explanations for a shift in bias scores, such as random error, or regression to the mean, could be responsible for changes in scores across time (Rodebaugh et al., 2016).

Consequently, other ways of analysing the data from the dot probe task have been recommended to try to improve both the task's test-retest reliability, and the ability to identify which cognitive processes are affected by successful bias modification. One such method of analysis that is in the early stages of application in this area is the diffusion decision model (Ratcliff, 1978). In this study, we apply this model to data from an attentional bias modification task to explore its ability to address issues of reliability, as well as its ability to identify cognitive process that are responsible for successful bias modification.

The Diffusion Decision Model

The diffusion decision model is based on the theory that when an individual chooses between two options, they accumulate evidence for one or the other option until a decision threshold is reached, and a response is executed. The model decomposes speeded two-choice decision making performance into four parameters: boundary separation, relative starting point, drift rate and non-decision time (see Figure 2 and Table 1). This study applied the diffusion decision model using the program *fast-dm 30* (Voss & Voss, 2007), in line with the prior research by Manuel et al. (under review).

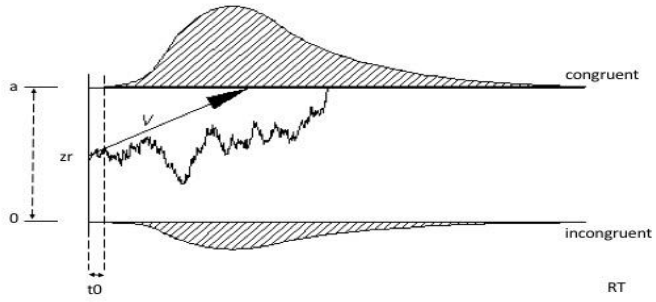


Figure 2. Diffusion decision model response parameters, boundary separation (*a*), relative starting point (z_r), drift rate (*v*), non-decision time (*t0*) for a decision where the response alternatives are congruent and incongruent, as per the dot probe task. (Diagram adapted from Voss & Voss, 2007; used with permission)

Boundary separation (*a*) reflects the participant's response caution. The two available choices are reflected by a lower decision boundary (0) and an upper decision boundary (*a*). Because the value of the lower boundary is fixed at 0, a larger value of *a* indicates a larger boundary separation, meaning that the individual requires more evidence to reach a decision (Voss & Voss, 2007). A larger boundary separation, holding all else equal, will result in responses that are on average slower but more accurate. This parameter is reflective of cognitive control, an important factor in response inhibition and emotion regulation.

Relative starting point (z_r) is the point between the decision boundaries from which evidence accumulation starts, measured as a proportion of boundary separation. This parameter reflects the bias an individual has for one response option over another based on their expectancy of that response derived from past experiences (White et al., 2016). The closer the starting point is to either boundary, the shorter the processing time will be for the corresponding response, and thus is indicative of a decisional bias. An unbiased starting point will be returned as 0.5, indicating an equal distance from either decision boundary.

Table 1.

Typical parameter values of the drift-diffusion model as returned by fast-dm.(Voss et al., 2015)

Parameter	Typical range	Description
Boundary separation (*a*)	$0.5 < a < 2$	Larger values equate to a more conservative decision style
Relative Starting Point (z_r)	$0.3 < zr < 0.7$	Deviations from 0.5 indicate that different amounts of evidence are required to make a decision for the alternative responses.
Drift rate (*v*)	$-5 < v < 5$	The rate of evidence accumulation. Values further from zero indicate stronger evidence. Positive values indicate evidence for the upper threshold response, negative values indicate evidence for the lower threshold response.
Non-decision time (*t0*)	$0.1 < t0 < 0.5$	Average duration of all non-decisional processes (encoding and response execution).

Drift rate (v) is the mean rate of evidence accumulation toward one choice or the other. Drift rate reflects the rate of approach to the decision threshold. A higher drift rate means faster accumulation of evidence. Drift rate is generally a function of stimulus or evidence quality. The easier an individual finds a stimulus to identify, the higher the drift rate. A higher drift rate for one emotion relative to others indicates a perceptual preference for that emotion.

Non-decision time (t0) is the time needed for sensory encoding prior to the evidence accumulation process, and for response execution once a decision boundary has been crossed. This parameter is the component of response time that is not related to decisional processes, and is

plausibly affected by hypervigilance or motor slowing, both of which are sometimes evident in individuals with anxiety and depression respectively.

Diffusion Analysis of Attentional and Interpretation Processes

Recently, Manuel, Kemps and McCarley, (under review; in preparation), used the diffusion decision model to analyse attentional and interpretation processes in a non-clinical sample. In one study, they analysed data returned from the dot probe task to explore the cognitive processes that underlie attentional bias. They hypothesised that attentional bias may be a function of the perceived strength of the stimulus that captures an individual's attention, as well as an implicit hypervigilant phasic response to that stimuli. Therefore, they examined the impact that presentation of different emotional stimuli had on the diffusion model parameters of drift rate and non-decision time in a non-clinical sample. Using facial stimuli that consisted of either a happy, sad, or angry expression paired with a neutral expression, they identified a tendency for evidence to accumulate more quickly for the existence of the probe when behind emotional faces rather than neutral faces. This was indicated by a slightly positive drift rate across all three emotional expressions. They did not find that evidence accumulated more quickly for one emotion over another. However, they did find that the parameters of the diffusion model were linked to clinical measures of anxiety and depression. Faster accumulation of evidence for angry stimuli was associated with depression and generalised anxiety, while slower non-decisional processes and more cautious responding were a protective factor to emotion regulation difficulties. Additionally, they found that a more conservative decision style was linked to better emotion regulation, and faster non-decisional processes were related to an increased difficulty regulating emotion.

In another study exploring the processes underlying interpretation bias, Manuel et al. (in preparation) used the diffusion decision model to analyse data from a yes/no task. As interpretation bias is primarily identified through judgments of ambiguous stimuli, they examined differences in cognitive processes that stem from identifying clearly identifiable emotional expressions, and milder, more ambiguous expressions. Because emotional expressions of a stronger intensity are

easier to identify than ambiguous expressions, a ceiling effect can mask differences between participants of different clinical status. In contrast, when the stimuli are more ambiguous, individuals who show an interpretation bias are hypothesised to identify the ambiguous stimuli as emotive more quickly than individuals who do not show an interpretation bias. Using a yes/no paradigm to explore the processes that underlie interpretation of ambiguous facial expressions relative to stronger intensity facial expressions, they identified that both perceptual (drift rate) and decisional (starting point) processes were affected. The participants presented with the ambiguous stimuli had lower drift rates and reduced starting point bias than the participants presented with the non-ambiguous stimuli. Thus, perceptual bias *and* decisional biases are affected by interpreting ambiguous expressions. Consequently, we can surmise that higher drift rates, and a starting point more strongly biased toward "yes" when presented with ambiguous information is an indicator of interpretation bias.

In our intervention study (see figure 2), we adopt the yes/no paradigm, using the mix of clear and ambiguous stimuli to identify any flow-on effects of attentional bias modification on the processes underlying interpretation bias. The yes/no paradigm is a simple computer task that requires participants to identify whether a single stimulus presented on a computer screen belongs to a target category. Stimuli that belong to the category in question are termed targets, and stimuli that do not belong to that category are termed distractors. The participant responds by pressing a corresponding key on the keyboard labelled "yes" or "no". When the stimulus presented is strongly associated with an attribute in line with the participant's implicit beliefs, response times for "yes" will be faster, indicating the presence of an interpretation bias.

Using an attentional bias task, the yes/no task, and an ABM training paradigm, we aim to identify how ABM impacts the cognitive processes that underlie attention and interpretation in individuals with a history of anxiety and depression. Further, we aim to provide insight into how the cognitive processes that are responsible for shifts in attentional and interpretation biases may differ between anxiety and depression, and individuals with no clinical history.

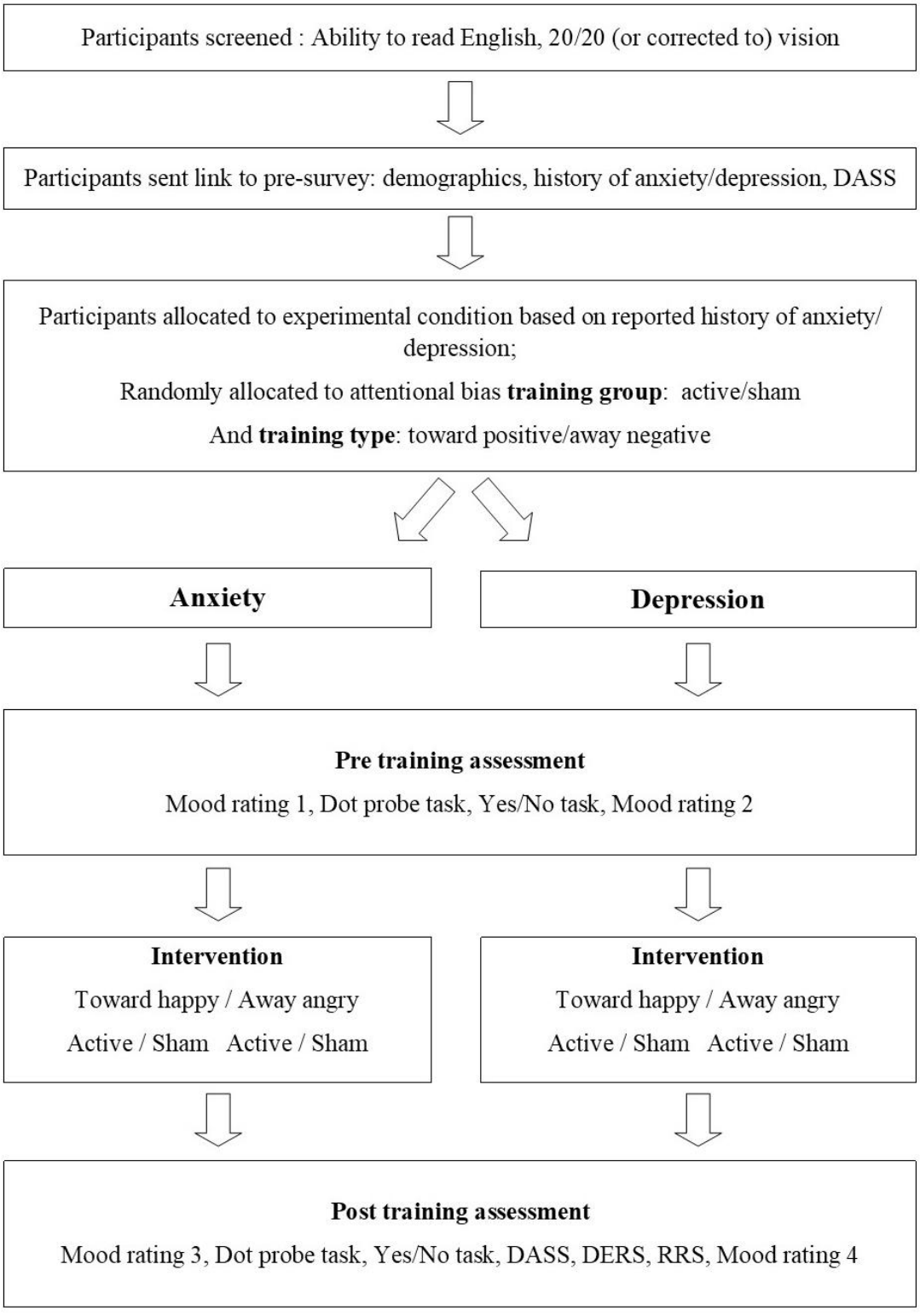


Figure 2. Training protocol for intervention study

We also explore links between these cognitive processes and scores on clinical measures of anxiety, depression, and emotion regulation difficulties. In addition, we explore whether training toward happy stimuli versus away from negative stimuli has a stronger impact on cognitive change, and ultimately, mood. Finally, we examine the reliability of diffusion model parameters to ascertain whether the diffusion model is a more reliable way to detect attentional bias than the current attentional bias score. We do this by assessing the test-retest reliability of the parameters of the diffusion model and comparing them with test-retest reliability of attentional bias scores.

To achieve these aims, we recruited individuals with a history of anxiety, depression, and individuals with no history of either disorder into our intervention study. We used an attentional bias modification paradigm to train individuals to attend toward happy stimuli, or to attend away from disorder-congruent negative stimuli (i.e., angry, sad). We also included a control condition for both the anxiety and depression cohorts who received sham training. It is the data from the control conditions that were used to assess stability of attentional bias scores and diffusion model parameters across time. We took measures of attentional and interpretation bias at pre-training to establish the existence of an attentional and/or interpretation bias and check for any group differences. We then administered the training paradigm, after which we took the pre- training measures again. The post-training scores were analysed against the pre-training scores.

Based on previous attentional bias research, we expect that individuals with a history of anxiety would have an attentional bias for angry faces, and individuals with a history of depression would have an attentional bias for sad faces. We hypothesise that this would be evident from attentional bias scores, as well as the diffusion model parameter of drift rate. We also hypothesise that there may be associated non-decisional differences due to hypervigilance or motor slowing, known symptoms of anxiety and depression, respectively (DSM-V; American Psychiatric Association, 2014). Further, we hypothesise that interpretation bias to disorder congruent stimuli would be evident from faster response times to “yes”, and higher drift rates and starting point values for angry and sad faces for individuals with a history of anxiety and depression, respectively. While

the investigation into the processes that change because of attentional bias training is exploratory, it is these same parameters that we expect to see shift over time in the groups who receive active training. We also expect that the groups that are trained toward happy stimuli may see an increase in mood relative to the groups who are trained to avoid negative stimuli.

Method

This study received ethics approval from the Flinders University Social and Behavioural Research Ethics Committee (Project number 6259: Investigating cognitive mechanisms underpinning emotional cognitive biases in individuals with varying levels of emotional disorders).

Participant Characteristics and Exclusion Criteria

Two hundred and seventy participants were recruited through the university's research participation program SONA. Eighty-four participants were undergraduate students who received course credit for a first-year psychology topic, and 186 participants received $20 remuneration. Eighteen participants were excluded from the final data due to incomplete data (8), performing at chance on the cognitive tasks (4), or disproportionate number of responses removed during data cleaning (6). Thus, data from 252 participants, were used in the final analysis. There were 199 females and 53 males with a mean age of 23 years (min 17, max 64, median 20). One hundred and twenty-eight participants identified as having no history of anxiety or depression, 66 reported a history of both anxiety and depression, 35 reported only anxiety, and 23 reported only depression.

Design

Two 6-factor mixed designs were used with between subjects' factors of disorder (anxiety, depression), training type (toward happy, away from negative) and training condition (active, sham), with the dot-probe and yes/no cognitive tasks containing the within subjects' factors. The dot probe task incorporated a 3 (emotion: happy, sad, angry) x 2 (probe: congruent, incongruent) x 2 (time: pre-training, post-training) repeated measures design, with dependent measures of RT and fast-dm parameters. The yes/no task incorporated a 4 (emotion: happy, sad, angry, neutral) x 2 (response: yes, no) x 2 (time: pre-training, post training) repeated measures design with dependent

measures of RT and fast-dm parameters. The dependent measures were also used to explore their relationships with scores on the clinical measures of depression, anxiety, rumination, and emotion regulation difficulties.

Materials

Stimuli. In line with other research in this area (Arndt & Fujiwara, 2012; Niles et al., 2013; Paulewicz et al., 2012; Tottenham et al., 2011), we selected facial stimuli from the NimStim database (Tottenham et al., 2009). This database consists of 646 facial expression stimuli across 43 faces of different gender and race. The database has expressions consisting of fearful, happy, sad, angry, surprised, calm, neutral, and disgusted, each with validity ratings (Tottenham et al., 2009). There are two sets of each expression. One set consists of the expressions with models' mouths open, and the other set consists of expressions with mouths closed. In line with the previous two studies in this thesis, faces with angry, happy, sad, and neutral expressions were used, and selected from the set with mouths closed. The mouths-closed stimuli were used to ease the morphing process that was used to create a mild subset of happy, sad and angry expressions. Three faces were excluded as the models did not present with all four (angry, happy, sad, neutral) expressions. The remaining stimuli consisted of digital photographs of 40 individuals (18 females and 22 males), each depicting each of the four expressions (happy, sad, angry, neutral). All pictures were presented in a black rectangular frame and cropped to just around the face to remove distinguishing external features, allowing the focus to be on the expression. All images were of equal size (332x458 pixels), resolution (120 pixels/inch), and RGB colour (8bpc). A subset of mild emotions was created by morphing each valence with its corresponding neutral face using the Fantamorph (Abrosoft, 2013) software program.

These mild faces were morphed to a strength of 50%, exactly halfway between the original stronger intensity emotion and the corresponding neutral face. This mild subset was incorporated into the yes-no task in line with previous research by Manuel et al. (in preparation).

Attentional bias assessment task: Dot probe task. (MacLeod & Mathews, 1998). The visual probe task consisted of 280 trials. For each trial, a centrally located fixation cross was displayed for 500ms, followed by the addition of a picture pair for another 500ms. Picture pairs comprised either a neutral and emotive (happy, sad, angry) facial expression of the same individual, or two of the same neutral facial expressions. Pictures were presented to the left and right of the fixation cross, centred horizontally, and equidistant from the centre of the screen. An arrow probe pointing up or down was presented in the location previously occupied by either the left or right image, centred horizontally. Presentation of the arrow was terminated by the participant pressing a corresponding directional response key. Response keys were the up and down arrow keys. A 500ms inter-stimulus interval followed the participant's response, after which the next trial began automatically.

The task was split into two blocks of 140 trials, allowing participants a self-determined break between blocks. A randomization algorithm determined stimulus presentation with constraints such that valence, trial type (congruent, incongruent, baseline), probe location (left, right) and probe direction (up, down) were counter-balanced within and across blocks, and that there were no more than three consecutive response presses on the one side to prevent response bias. There were 80 trials per valence (happy, sad, angry), with 40 congruent (probe behind valence) and 40 incongruent (probe behind neutral) trials, as well as 40 neutral/baseline (both pictures neutral) trials.

Interpretation bias: Yes/No task. (Manuel et al., in preparation). The Yes/No task consisted of four blocks of 80 trials. For each trial, a single face appeared on a black background centred horizontally and vertically, and remained until the participant made a response indicating that the expression did or did not match a target emotion. Participants were instructed to respond using the keyboard, as quickly and accurately as possible. Response keys were the home key for "yes" and the end key for "no", labelled accordingly. A 500ms inter-stimulus interval followed the participant's response, after which the next trial began automatically. At the beginning of each

block participants were asked, “Are the following faces neutral/happy/sad/angry?” The stimulus face matched the target valence on 2/3 (56 trials) of the trials within each block, to encourage the “yes” bias. On remaining trials (24 trials), the stimulus faces expressed one of the 3 distractor emotions for that block, for example, if happy was the target emotion, distractor emotions were angry, sad and neutral. Target and distractor stimuli were generated randomly using a randomization algorithm, with mild and stronger intensity expressions generated equally. The order of the blocks was counterbalanced between participants. Stimuli consisted of a mix of milder, more ambiguous facial expressions, as well as clearly identifiable facial expressions.

Training paradigm. The dot probe training paradigm used the same task and instructions as the dot probe task. However, in the training paradigm, there were 320 trials in total, split into two blocks of 160 trials. Participants were presented with either happy and neutral stimulus pairs (training toward happy condition) or neutral and angry/sad stimulus pairs (training away from negative condition). In the active training conditions, the probe was behind the happy (training toward happy condition) or neutral (training away from negative condition) expression for 90% of trials, and appeared behind the neutral (training toward happy condition) or angry/sad (training away from negative condition) expression for the remaining trials. In the sham training conditions, the probe appeared behind each expression an equal 50% of trials. The choice to use a 90/10 training ratio was used, as opposed to a 100/0 ratio, to ensure participants maintained concentration and to reduce the obviousness of the contingency (Schoenmakers, Wiers, Jones, Bruce, & Jansen, 2007).

Clinical measures.

Difficulties with emotion regulation scale (DERS; Gratz & Roemer, 2004). This 36-item self-report questionnaire consists of six dimensions of emotion regulation: (a) Non-acceptance of emotional reactions; (b) Difficulty engaging in goal-directed behaviour; (c) Difficulty controlling impulses; (d) Lack of emotional awareness; (e) Limited access to emotion-regulation strategies; and (f) Lack of emotional clarity. The DERS begins each item with the phrase, “When I’m upset…”.

Participants rate each item on a five-point Likert scale from “almost never” (1) to “almost always” (5). Eleven items are reverse scored, and the total score can range from 36 to 180 with a higher score representing increased difficulty with emotion regulation. In non-clinical samples of adults, the DERS has an average total score of 75-80, there are no specific established norms for clinical groups, however, scores above this average are considered indicative of emotion regulation difficulties (Gratz & Roemer, 2004b). Both the total score and the difficulty controlling impulses subscale were used in this study (hereafter referred to as the impulsiveness subscale). The DERS has high internal consistency (α =.93), good test-retest reliability (ρ_I=.88), and satisfactory construct and predictive validity (Gratz & Roemer, 2004). Within this sample, Cronbach’s alpha was .95.

Depression Anxiety and Stress Scale short form (DASS-21; Lovibond & Lovibond, 1996). This self-report questionnaire consists of 21 items that divide into three subscales, each consisting of 7 four-point Likert rating-scales measuring depression, anxiety and stress. The questionnaire asks the respondent to rate how much each statement applied to them over the past week. The rating scale ranges from 0 (Did not apply to me at all) to 3 (Applied to me very much, or most of the time). The DASS-21 has high internal consistency (α ranging from .82-.93; Antony et al., 1998; Henry & Crawford, 2005) and moderate levels of concurrent validity (*r* =.46-.85; Antony et al., 1998). Within the current sample, alpha was .93. The depression subscale was used to measure depressive symptoms. The anxiety subscale loads onto diagnostic criteria for various anxiety disorders *except* generalised anxiety disorder (GAD), whereas the stress subscale loads onto diagnostic criteria for generalised anxiety disorder (Lovibond & Lovibond, 1995). As we are interested in generalised anxiety, we used the stress subscale in this study. Scores range from 0-28+ for depression, and 0-32+ for stress, with higher scores indicating increased level of symptom severity.

Visual analogue mood scale. (VAMS; Ahearn, 1997). This measure is a single line, typically 10cm long), ordinal scale that is anchored by two statements or images reflecting opposite extremes of a single mood construct. For example, in this study, the VAMS was presented on the

computer screen asking participants to rate their mood. Positioned to the left-hand side of the line was the depiction of a simple sad-face emoticon, whereas to the right-hand side of the line was a simple happy-face emoticon. Participants were required to identify their current mood by clicking on the line. Score range from 0-100, with higher score indicative of more positive mood.

Procedure

At the time of signing up to the study, participants completed a pre-survey, which took approximately 5 minutes. This screening survey asked for demographic information, included the DASS-21 screening tool, and asked if the participants had a history of anxiety, depression, both, or neither. Based on this answer alone, participants were then randomly allocated to a training condition. Individuals with a history of anxiety were randomly allocated to either the active or sham training condition for the anxiety group, likewise, individuals with a history of depression were randomly allocated to the sham or active training condition in the depression group. Participants who reported a history of both, or no history at all were randomly allocated to group and condition. The DASS-21 from the screening survey was only used in final analysis to compare the groups.

When arriving at the lab for their allocated time slot, all other tasks were completed in one session that lasted approximately 80 minutes. Each task was displayed on a computer screen. A visual analogue mood scale was used to capture snapshots of mood throughout the session, to assess changes in mood between groups. The visual analogue mood scale was presented four times. It was presented first, followed by the dot-probe and yes/no measurement tasks to obtain pre-training scores. The order of these first assessment tasks was counterbalanced between participants. The visual analogue mood scale was presented again, followed by the training task and another mood scale. The dot probe and yes/no tasks, counter balanced between participants, were again administered to obtain post training scores. A final visual analogue mood scale was presented again, followed by the clinical measures. Clinical measures were presented in Qualtrics, an online survey program. They were presented in the same order for each participant with the DASS presented first, followed by the DERS.

Analysis

Data cleaning and preparation.

Dot Probe task. As per MacLeod and colleagues (1986), all error responses were removed from analysis of RT data from the pre- and post-training assessment tasks. On average, accuracy rates were 96.26% at pre-training and 94.08% at post-training. For each trial type (congruent/incongruent) within each valence (happy, sad, angry, neutral), RTs that were 3 or more standard deviations above each subject's mean RT were removed, as were responses less than 150ms. This resulted in an additional 1.41% of responses being removed from the pre-training data set, and 1.85% from the post-training data set. For each valence, an overall bias score for each participant was calculated by taking the mean RT for congruent trials and subtracting it from the mean RT for incongruent trials. Thus, positive scores indicate an attentional bias toward the emotional stimuli, whereas negative scores indicate a bias toward neutral stimuli and away from the emotional stimuli.

Fast-dm analysis. Data were cleaned in the same way as for the traditional analysis with the exception that errors were not excluded. The data were coded such that congruent responses were mapped to the upper response threshold and incongruent responses were mapped to the lower threshold, and neutral trials were removed. Fast-dm 30 was run in line with recommendations from Voss, Voss, and Lerche's (2015) tutorial paper. The fast-dm 30 control file was set to allow drift rate and non-decision time to vary based on emotion across participants, and starting point and boundary separation were free to vary across participants but not across emotions. Additional parameters of variability in starting point, drift rate, and non-decision time can also be introduced into model analysis. While more parsimonious models can be preferable to more complex models (Lerche & Voss, 2016), it has been recommended to include non-decision time variability in the model to achieve stable parameter estimates (Voss et al., 2015; Lerche and Voss, 2016). Therefore, inter-trial variability of non-decision time (s_{t0}) was also allowed to vary across participants.

Model fit was assessed by the Kolmogorov–Smirnow (KS) statistic provided by the fast-dm 30 software. The KS statistic assesses whether the distribution of observed data differs significantly from a predicted distribution. The KS fit statistic returned by fast-dm 30 did not reveal any significant deviations between empirical and estimated RT distributions for any of the participants. As Voss et al. (2015) caution against relying solely on the fit statistic returned by fast-dm 30 and recommend plotting the data graphically. Comparison of empirical and predicted response time distributions for each expression (happy, angry, sad, neutral) were plotted, along with quantile probability plots comparing empirical and predicted response times for the 25th, 50th, and 75th percentiles of each expression (happy, sad, angry). These plots are displayed in Appendix 3.

Yes/No task. On average, participant accuracy rates were 87.82% at pre-training and 84.61% at post-training. RTs less than 150ms and more than 3 standard deviations from each participant's mean RT for each response (yes, no) within each emotional condition (happy, sad, angry, neutral) and by strength (strong, mild, neutral) were removed. This accounted for 1.45% of responses for both pre-training and post-training data sets.

Fast-dm analysis. Cleaning was as per traditional analysis. Responses were coded as per their button responses: "Yes" responses for the upper threshold, "no" responses for the lower threshold. Fast-dm was run separately for each emotion (Voss, Voss & Lerche, 2015), and the fast-dm 30 control file was set to allow drift rate to vary based on stimulus type (yes strong, yes ambiguous, and no – the average of which was computed and used in the final drift rate analysis), starting point, boundary separation, non-decision time, and non-decision time variability to vary across participants. As with the dot-probe analyses, model fit was assessed by the Kolmogorov–Smirnow (KS) statistic provided by the fast-dm 30 software. The KS fit statistic returned by fast-dm 30 did not reveal any significant deviations between empirical and estimated RT distributions for any of the participants. Similarly, comparison of empirical and predicted response time distributions for each expression (happy, angry, sad, neutral) were plotted, along with quantile probability plots

comparing empirical and predicted response times for the 25th, 50th, and 75th percentiles of each expression (happy, sad, angry). These plots are displayed in Appendix 3.

DASS, DERS & RSS. Clinical measures were checked for skewness and kurtosis, and were within acceptable limits (+/- 2). There were no outliers.

Examination of group differences pre-training

Cognitive measures. Mixed ANOVAs were performed to analyse differences between response times and fast-dm parameters at pre-training for individuals with self-reported history of anxiety or depression and individuals with no such history. Data were analysed separately for anxiety and depression. Bonferroni correction was applied, and where sphericity tests indicated it applicable, the Greenhouse Gausser correction was applied.

Relationships between cognitive and clinical measures. To assess the relationships between attentional bias and interpretation bias, and the clinical measures correlations were used. We examined the relationships between traditional bias scores and fast-dm parameters for both tasks at pre-training and scores on the clinical outcome measures of depression/stress and the DERS.

Training effects.

Cognitive measures. Data were analysed separately for anxiety and depression. Mixed ANOVAs were performed to analyse training differences between training condition (active, sham), training type (toward happy, away sad/angry), and history (history, no history) across time (pre-training, post-training). For traditional dot-probe analysis we examined these group effects across time, emotion (happy, sad, angry) and response type (congruent/incongruent). For the fast-dm parameters of drift rate and non-decision time derived from the dot-probe data, we examined the group differences across time and emotion (happy, sad, angry), whereas for start-point and boundary separation, we analysed the group differences across time. For RT analysis of the interpretation bias task we also examine these group differences across time, emotion (happy, sad, angry, neutral), and response (yes, no). For the fast-dm parameters, we analysed the group

differences across time and emotion (happy, sad, angry, neutral). Bonferroni correction was applied, and where sphericity tests indicated it applicable, the Greenhouse Gausser correction was applied.

Relationships between cognitive and clinical measures. Correlations were carried out to identify whether changes in bias scores or fast-dm parameters were related to scores on the clinical outcome measures of anxiety/depression and the DERS. We also explored the impact of training on mood ratings using a mixed ANOVA comparing the impact of training, training type, history and time.

Reliability analysis. To assess reliability of the attentional bias measure and the fast-dm measures, correlations were conducted on the pre-test and post-test scores for participants in the sham training conditions.

Results

Group characteristics prior to assessment and training are presented in Table 2. The mean age was similar across groups, and the clinical scores for both anxiety, $t(124) = 6.24$, $p = .001$, and depression, $t(124) = 5.06$, $p < .001$, were significantly higher for individuals with a self-reported history than those without such a history.

Table 2

Pre- study group characteristics by reported history of anxiety and depression.

	Anxiety				Depression			
	History		No history		History		No History	
	n = 63		n=63		n=61		n=65	
Age	22.94	(7.48)	22.49	(8.31)	24.21	(8.48)	22.15	(6.02)
M/F	8	55	15	48	14	47	16	49
Clinical score*	9.57	(3.75)	5.38	(3.78)	8.77	(5.09)	4.71	(3.89)

*Clinical score is the Stress subscale for the anxiety condition (corresponds to generalised anxiety disorder), and the Depression subscale for the depression condition.

Pre-training measures

To be able to assess the efficacy of an attentional bias modification study on attentional and interpretation biases, we first needed to establish that attentional and interpretation biases were captured in the current sample. We examined RTs and diffusion model parameters to identify if there was evidence of these biases in the current sample.

Attentional bias.

Anxiety.

Response times. Correct trial mean RTs for congruent and incongruent trials for each emotion (happy, sad, angry) of the dot probe task are plotted for each group (history/no history) in Figure 2.

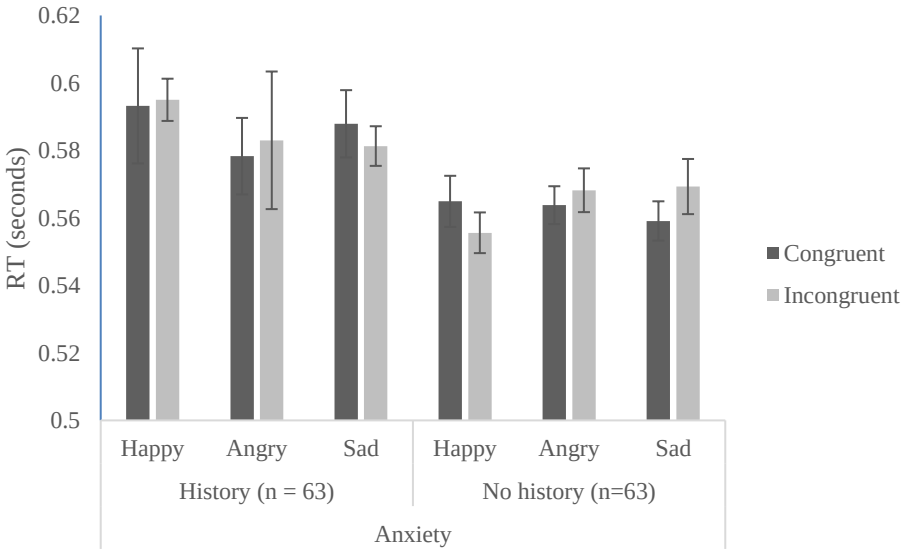


Figure 2. Dot probe RTs (seconds) by trial type and history for correct trials. Congruent trials = probe behind emotional face; incongruent trials = probe behind neutral face. Error bars represent within-subjects 95% confidence intervals[4].

[4] All confidence intervals throughout this paper were calculated using the Cousineau-Morey method (Morey, 2008).

Attentional Bias. A 2 (group: history, no history) x 2 (trial type: congruent, incongruent) x 3 (emotion: happy, sad, angry) mixed ANOVA did not identify an interaction between RTs for group, trial, and emotion, $F(1.54,191.27) = 2.78$, $p = .08$, $\eta^2_{partial} = .02$, indicating no evidence for an attentional bias that differed by history of anxiety. Additionally, there was no attentional bias across the sample, $F(1.54,191.27) = 1.02$, $p = .35$, $\eta^2_{partial} = .01$, and RTs did not differ by group, $F(1,124) = 1.16$, $p = .28$, $\eta^2_{partial} = .01$.

Further, for the group with a history of anxiety, attentional bias scores did not differ from zero for any emotion (*p*s >.05). In contrast, the group with no clinical history showed a small attentional bias for happy stimuli, $t(62) = 2.22$, $p = .03$, $d = 0.27$, but not for angry or sad stimuli (*p*s >.05). Therefore, our hypothesis that we would find an attentional bias for angry expressions in the group of participants with a history of anxiety was not supported.

Diffusion Model parameters. To assess whether there was a group difference identified by the diffusion model parameters, we conducted two 2 (group: history, no history) x 3 (emotion: happy, sad, angry) mixed ANOVAs for drift rate and non-decision time. To examine the differences between groups for starting point and boundary separation, we ran independent samples t-tests.

Drift rate. There was no interaction between group and emotion on drift rate, $F(2, 248) = 0.76$, $p = .47$, $\eta^2_{partial} = .006$, indicating that our hypothesis that drift rate would be higher for the probe behind angry stimuli for individuals with a history of anxiety was also not supported. One sample *t*-tests showed that drift rate was positive for all three emotions in the group with a self-reported history of anxiety (happy $t(62) = 2.99$, $p = .004$, $d = 0.36$; sad $t(62) = 2.64$, $p = .01$, $d = 0.35$, angry $t(62) = 3.29$, $p = .002$, $d = 0.41$), indicating that evidence accumulated slightly faster for the probe on congruent trials than on incongruent trials. For individuals with no history of anxiety, rate of evidence accumulation for the probe on congruent versus incongruent trials did not differ (*p* values > .05), but the difference in drift rate between groups was not significant, $F(1, 124) = 3.34$, $p = .07$, $\eta^2_{partial} = .03$.

Non-decision time. There was no interaction between group and emotion, $F(2, 248) = 0.01$, $p = .99$, $\eta^2_{partial} < .001$, indicating no differences in non-decision time across emotions between individuals with a history of anxiety and individuals without such a history. Nor did non-decision time differ between group $F(1, 124) = 0.09$, $p = .77$, $\eta^2_{partial} = .001$.

Starting point and boundary separation. There was no difference between groups for starting point, $t(124)=0.79$, $p = .43$, $d < 0.001$, or boundary separation, $t(124)=0.001$, p 1.00, $d < 0.001$, which is what we had anticipated.

Depression.

Response times. Correct trial mean RTs for congruent and incongruent trials for each emotion (happy, sad, angry) of the dot probe task are plotted for each group (history/no history) in Figure 3.

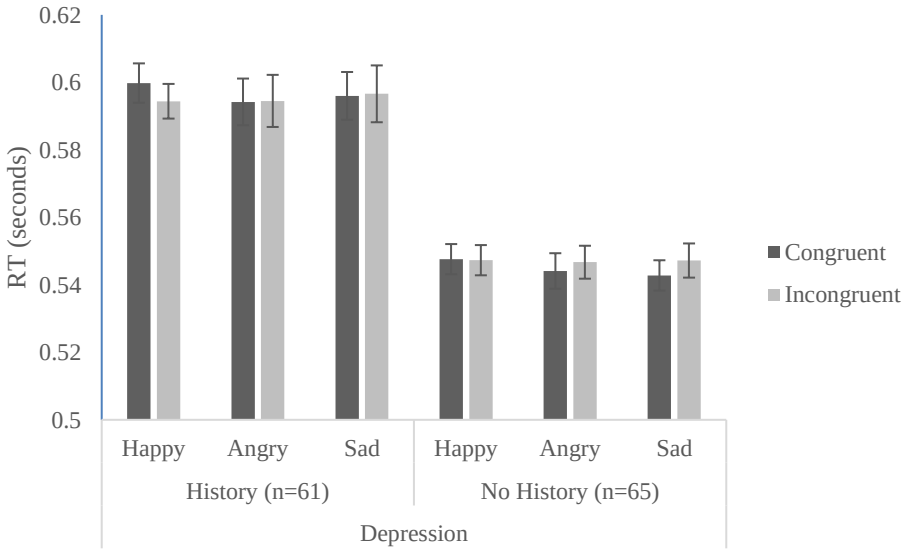


Figure 3. Dot probe RTs (seconds) by trial type and history for correct trials. Congruent trials = probe behind emotional face; incongruent trials = probe behind neutral face. Error bars represent within-subjects 95% confidence intervals.

Attentional Bias. A 2 (group: history, no history) x 2 (trial type: congruent, incongruent) x 3 (emotion: happy, sad, angry) mixed ANOVA found no interaction between RTs for history, emotion, and trial type, $F(1.83, 226.42) = 0.06$, $p = .93$, $\eta^2_{partial} = .001$, indicating no evidence for an

attentional bias that differed by history of depression. There was no attentional bias across the sample, $F(1.83, 226.42) = 1.03$, $p = .35$, $\eta^2_{partial} = .01$, and no attentional bias score differed from zero (*p*s >.05). However mean RTs were 50ms slower for the group with a history of depression than the group with no such history, $F(1,124) = 11.53$, $p = .001$, $\eta^2_{partial} = .09$. Therefore, our hypothesis that we would find an attentional bias for sad expressions in the group of participants with a history of depression was not supported.

Diffusion Model parameters. To assess whether there was a group difference identified by the diffusion model parameters, we conducted two 2 (group: history, no history) x 3 (emotion: happy, sad, angry) mixed ANOVAs for drift rate and non-decision time. To examine the differences between groups for starting point and boundary separation, we ran independent samples t-tests.

Drift rate. There was no interaction between emotion and group, $F(2, 248) = 2.40$, $p = .09$, $\eta^2_{partial} = .02$ on drift rate (*v*), indicating no difference in drift rate between groups based on emotion. However, there was a main effect of group, $F(1, 124) = 6.17$, $p = .01$, $\eta^2_{partial} = .05$, indicating that drift rate was higher for individuals with a self-reported history of depression than individuals without this history. One sample *t*-tests for each group indicated that for individuals with a history of depression, drift rate was positive for all three emotions (happy $t(60)=3.47$, $p = .001$, $d = 0.46$, sad $t(60)=3.38$, $p = .001$, $d = 0.44$, angry, $t(60) = 5.05$, $p < .001$, $d = 0.67$), but for individuals with no history of depression, drift rate was only positive for sad facial expressions, $t(60)=2.12$, $p = .04$, $d = 0.28$. Therefore, drift rate for individuals with a history of depression was faster for congruent than incongruent trials across all emotions, suggesting that attention was captured by emotional stimuli in general. Thus, our hypothesis that individuals with a history of depression would have a stronger processing bias for sad stimuli was not supported.

Non-decision time. There was no interaction between group and emotion, $F(1.87, 232.18) = 1.05$, $p = .35$, $\eta^2_{partial} < .008$, indicating no differences in non-decision time across emotions between individuals with a history of depression and individuals without such a history. The group with a history of depression had a slightly faster non-decision time than the group with no history,

$F(1, 124) = 7.28$, $p = .008$, $\eta^2_{partial} = .06$, suggesting that rather than motor slowing, individuals with a history of depression have a more hypervigilant response than individuals with no history of depression.

Starting point and boundary separation. There was no difference in starting point between individuals with a history of depression and individuals without this history, $t(124)=1.89$, $p = .06$, $d = 0.13$, nor was there a difference between groups for boundary separation, $t(124) = 0.25$, $p = .80$, $d < 0.001$. Taken all together, the analysis of the dot-probe task did not capture a disorder-congruent bias for anxiety or depression by RTs or diffusion parameters.

Interpretation bias

Anxiety

Response times. Response times for yes and no responses are displayed in Figure 5. A 4 (emotion: happy, sad, angry, neutral) x 2 (response: yes, no) x 2 (group: history, no history) mixed ANOVA found no interaction for RTs between history, emotion and trial type, $F(2.82, 349.10) = 2.59$, $p = .06$, $\eta^2_{partial} = .02$, indicating no difference in interpretation bias between groups. There was an interaction between emotion and trial type, $F(2.82, 349.10) = 5.36$, $p = .002$, $\eta^2_{partial} = .04$. A post hoc repeated measures ANOVA revealed that RTs for yes responses differed by emotion, $F(2.70, 337.18) = 11.16$, $p < .001$, $\eta^2_{partial} = .08$, with yes responses to happy stimuli the fastest.

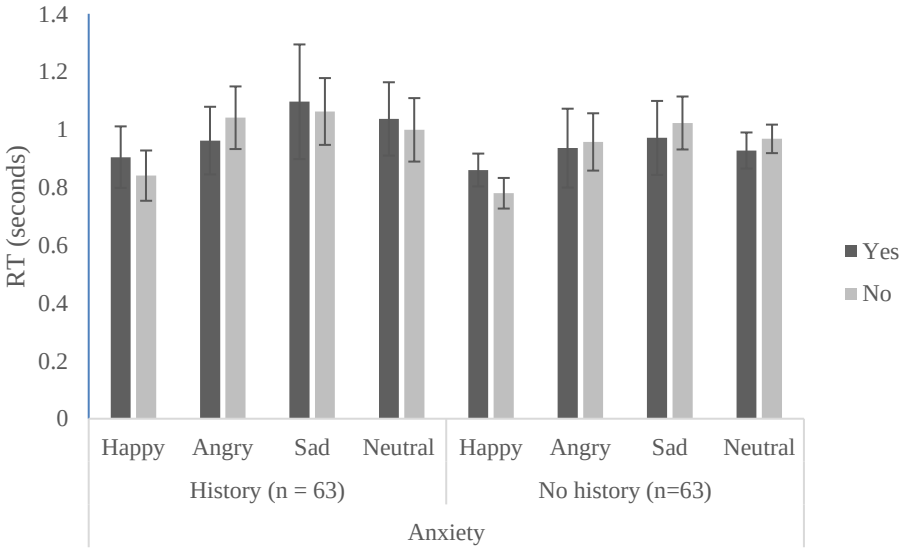


Figure 5. Mean response times (ms) for each response (yes, no) and emotion (neutral, sad, happy, angry) for individuals with a self-reported history of anxiety and individuals without; error bars represent within-subjects 95% confidence intervals.

Diffusion parameters. To assess whether there was a group difference identified by the diffusion model parameters, we ran a series of 4 (emotion: happy, sad, angry, neutral*)* x 2 (group: history, no history) mixed ANOVAs.

Boundary separation (a). There was no interaction between emotion and group, $F(3, 372) = 0.18$, $p = .91$, $\eta^2_{partial} = .001$, and there was no main effect of group, $F(1, 124) = 2.54$, $p = .11$, $\eta^2_{partial} = .02$. Therefore, our hypothesis that boundary separation would be smaller for the angry condition for individuals with a history of anxiety than those with no such history was not supported.

Relative starting point. There was no interaction between emotion and group, $F(2.75, 341.30) = 0.12$, $p = .94$, $\eta^2_{partial} = .001$, or main effect of group, $F(1,124) = 0.71$, $p = .79$, $\eta^2_{partial} = .001$, indicating that relative starting point did not differ between groups. We had predicted that starting point would be higher in the angry condition for individuals with a history of anxiety than individuals without this history. However, this was not supported.

Drift rate (v). There was no interaction between emotion and group, $F(2.58, 319.79) = 1.83$, $p = .15$, $\eta^2_{partial} = .01$; or group, $F(1,124) = 0.60$, $p = .44$, $\eta^2_{partial} = .01$, indicating no differences in the rate of evidence accumulation for individuals with a self-reported history of anxiety than for individuals with no such history. We had expected that individuals with a history of anxiety would have accumulated evidence more quickly for angry stimuli, and thus our hypothesis was not supported.

Non-decision time (t0). There was no interaction between emotion and group, $F(3, 372) = 0.44$, $p = .72$, $\eta^2_{partial} = .004$, nor was there a main effect of group, $F(1, 124) = 0.38$, $p = .57$, $\eta^2_{partial} = .003$. This did not support our hypothesis of a hypervigilant response in the group with a history of anxiety.

Depression

Response times. Response times for yes and no responses are displayed in Figure 6. A 4 (emotion: happy, sad, angry, neutral) x 2 (response: yes, no) x 2 (group: history, no history) mixed ANOVA found no interaction between RTs for history, emotion and trial type, $F(2.74, 339.45) = 1.35$, $p = .23$, $\eta^2_{partial} = .01$, indicating no difference in interpretation bias between groups. There was a main effect of group $F(1, 124) = 22.31$, $p = .001$, $\eta^2_{partial} = .09$, such that mean response times were on average 130ms slower in the group with a self-reported history of depression than the group with no such history. However, our hypothesis that "yes" responses in the sad condition would be faster for the group with a history of depression than the group without this history, was not supported.

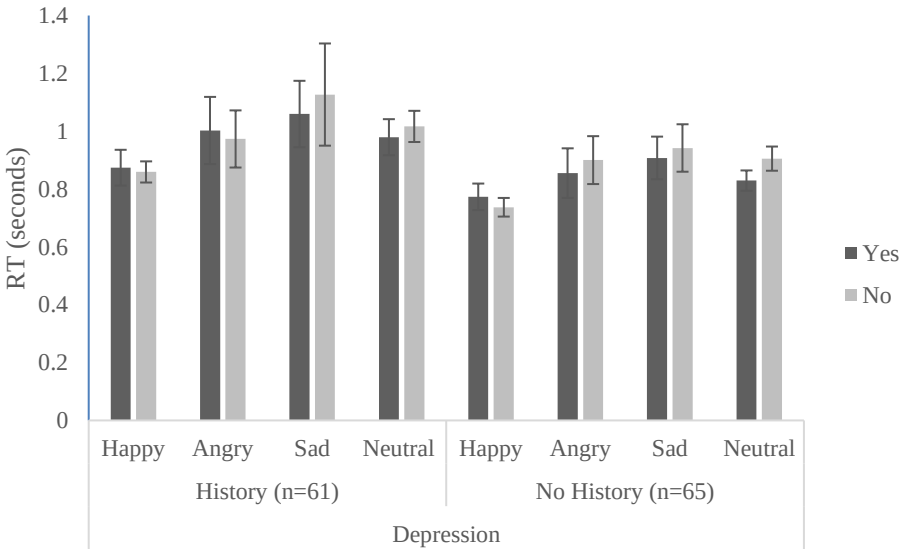


Figure 6. Mean response times (ms) for each response (yes, no) and emotion (neutral, sad, happy, angry) for individuals with a self-reported history of depression and individuals without; error bars represent within-subjects 95% confidence intervals.

Diffusion parameters. To assess whether there was a group difference identified by the diffusion model parameters, we ran a series of 4 (emotion: happy, sad, angry, neutral*)* x 2 (group: history, no history) mixed ANOVAs.

Boundary separation (a). There was no interaction between emotion and stimulus type, $F(3, 372) = 0.21$, $p = .89$, $\eta^2_{partial} = .002$; however, there was a main effect of group, $F(1, 124) = 9.39$, $p = .003$, $\eta^2_{partial} = .07$, indicating a more conservative response style for the group with a history of depression than the group with no such history.

Relative starting point (zr). The interaction between emotion and group was not significant, $F(3,372) = 1.38$, $p = .25$, $\eta^2_{partial} = .01$, and there was no main effect of group, $F(1, 124) = 2.23$, $p = .14$, $\eta^2_{partial} = .02$. These findings did not support our hypothesis that individuals with depression would have a higher starting point for sad expressions than individuals with no such history.

Drift rate (v). There was no interaction between emotion and group, $F(2.78, 345.08) = 0.45$, $p = .72$, $\eta^2_{partial} = .004$, nor was there a main effect of group, $F(1,124) = 0.35$, $p = .56$, $\eta^2_{partial} = .003$, on drift rate. Therefore, there was no evidence to support our hypothesis that individuals with a history of depression would have a faster drift rate for sad stimuli than individuals without a history of depression.

Non-decision time (t0). There was no interaction between emotion and stimulus type, $F(2.81, 348.34) = 0.67$, $p = .56$, $\eta^2_{partial} = .005$, nor was there a main effect of group, $F(1, 124) = 1.42$, $p = .24$, $\eta^2_{partial} = .01$; thus, there was no indication of motor slowing in the group with a history of depression. As with the dot probe task, the results of the interpretation bias analyses revealed that there was no disorder-congruent interpretation bias captured in the current sample.

Training Effects

Attentional bias. Descriptive statistics for the anxiety cohort are presented in Tables 3 and 4. Table 3 reports the mean RTs and fast-dm 30 parameter values and SDs for training away from angry at pre- and post-training; Table 4 reports the mean RTs and fast-dm 30 parameter values and SDs for training toward happy at pre- and post-training. Similarly, descriptive statistics for the

depression cohort are presented in Tables 5 and 6. Table 5 reports the mean RTs and fast-dm 30 parameter values and SDs for training away from sad at pre- and post-training; Table 6 reports the mean RTs and fast-dm 30 parameter values and SDs for training toward happy at pre- and post-training.

Table 3

Pre- and post-training descriptive statistics for the dot probe for individuals in the anxiety cohort who were given the "away from angry" training paradigm

			Active								Sham							
			History n=15				No History n=15				History n=15				No History n=16			
			Pre		Post		Pre		Post		Pre		Post		Pre		Post	
			M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RTs																		
	Happy	Con	0.57	0.09	0.58	0.08	0.54	0.08	0.52	0.05	0.55	0.05	0.56	0.06	0.55	0.09	0.54	0.08
		Inc	0.56	0.09	0.59	0.08	0.54	0.07	0.53	0.06	0.56	0.05	0.56	0.05	0.56	0.09	0.54	0.08
	Sad	Con	0.56	0.08	0.57	0.08	0.54	0.07	0.54	0.06	0.56	0.05	0.55	0.05	0.55	0.09	0.54	0.07
		Inc	0.56	0.08	0.59	0.08	0.54	0.07	0.52	0.06	0.56	0.05	0.57	0.06	0.55	0.07	0.54	0.09
	Angry	Con	0.57	0.09	0.58	0.07	0.54	0.07	0.54	0.07	0.56	0.05	0.57	0.05	0.56	0.07	0.55	0.09
		Inc	0.57	0.10	0.57	0.06	0.53	0.07	0.52	0.05	0.56	0.06	0.56	0.05	0.55	0.09	0.54	0.08
Fast-dm parameters																		
a			0.71	0.12	0.76	0.12	0.70	0.15	0.73	0.11	0.72	0.13	0.81	0.20	0.73	0.10	0.80	0.14
z_r			0.48	0.03	0.49	0.03	0.48	0.02	0.48	0.03	0.48	0.03	0.49	0.04	0.49	0.02	0.48	0.03
v	Happy		0.15	0.22	-0.01	0.19	0.10	0.20	0.06	0.17	0.11	0.27	0.01	0.28	-0.01	0.20	-0.01	0.24
	Sad		0.07	0.27	0.09	0.20	0.01	0.13	0.05	0.15	0.09	0.22	0.03	0.27	0.12	0.22	-0.01	0.24
	Angry		0.15	0.30	0.01	0.25	0.13	0.14	0.09	0.27	0.06	0.24	-0.01	0.20	0.04	0.24	0.14	0.27
t0	Happy		0.42	0.05	0.41	0.06	0.43	0.05	0.40	0.04	0.47	0.09	0.44	0.07	0.44	0.09	0.41	0.06
	Sad		0.43	0.05	0.41	0.07	0.43	0.05	0.40	0.05	0.46	0.09	0.43	0.08	0.44	0.08	0.40	0.07
	Angry		0.43	0.04	0.41	0.07	0.43	0.05	0.41	0.05	0.47	0.09	0.44	0.07	0.45	0.08	0.40	0.07

Table 4

Pre- and post-training descriptive statistics for the dot probe for individuals in the anxiety cohort who were given the "toward happy" training paradigm

			Active								Sham							
			History n=16				No History n=16				History n=15				No History n=16			
			Pre		Post		Pre		Post		Pre		Post		Pre		Post	
			M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RTs																		
	Happy	Con	0.59	0.11	0.58	0.08	0.54	0.08	0.52	0.05	0.55	0.05	0.56	0.06	0.55	0.09	0.54	0.08
		Inc	0.59	0.11	0.59	0.08	0.54	0.07	0.53	0.06	0.56	0.05	0.56	0.05	0.56	0.09	0.54	0.08
	Sad	Con	0.58	0.11	0.57	0.08	0.54	0.07	0.54	0.06	0.56	0.05	0.55	0.05	0.55	0.09	0.54	0.07
		Inc	0.59	0.11	0.59	0.08	0.54	0.07	0.52	0.06	0.56	0.05	0.57	0.06	0.55	0.07	0.54	0.09
	Angry	Con	0.59	0.11	0.58	0.07	0.54	0.07	0.54	0.07	0.56	0.05	0.57	0.05	0.56	0.07	0.55	0.09
		Inc	0.59	0.12	0.57	0.06	0.53	0.07	0.52	0.05	0.56	0.06	0.56	0.05	0.55	0.09	0.54	0.08
Fast-dm parameters																		
a			0.72	0.13	0.76	0.12	0.70	0.15	0.73	0.11	0.72	0.13	0.81	0.20	0.73	0.10	0.80	0.14
z_r			0.47	0.02	0.49	0.03	0.48	0.02	0.48	0.03	0.48	0.03	0.49	0.04	0.49	0.02	0.48	0.03
v	Happy		0.11	0.18	-0.01	0.19	0.10	0.20	0.06	0.17	0.11	0.27	0.01	0.28	-0.01	0.20	-0.01	0.24
	Sad		0.08	0.12	0.09	0.20	0.01	0.13	0.05	0.15	0.09	0.22	0.03	0.27	0.12	0.22	-0.01	0.24
	Angry		0.10	0.14	0.01	0.25	0.13	0.14	0.09	0.27	0.06	0.24	-0.01	0.20	0.04	0.24	0.14	0.27
t0	Happy		0.43	0.07	0.41	0.06	0.43	0.05	0.40	0.04	0.47	0.09	0.44	0.07	0.44	0.09	0.41	0.06
	Sad		0.42	0.06	0.41	0.07	0.43	0.05	0.40	0.05	0.46	0.09	0.43	0.08	0.44	0.08	0.40	0.07
	Angry		0.43	0.06	0.41	0.07	0.43	0.05	0.41	0.05	0.47	0.09	0.44	0.07	0.45	0.08	0.40	0.07

Table 5

Pre- and post-training descriptive statistics for the dot probe for individuals in the depression cohort who were given the "away from sad" training paradigm

			Active								Sham							
			History n=16				No History n=17				History n=15				No History n=16			
			Pre		Post		Pre		Post		Pre		Post		Pre		Post	
			M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RTs	Happy	Con	0.61	0.12	0.58	0.08	0.54	0.08	0.52	0.05	0.55	0.05	0.56	0.06	0.55	0.09	0.54	0.08
		Inc	0.60	0.12	0.59	0.08	0.54	0.07	0.53	0.06	0.56	0.05	0.56	0.05	0.56	0.09	0.54	0.08
	Sad	Con	0.61	0.13	0.57	0.08	0.54	0.07	0.54	0.06	0.56	0.05	0.55	0.05	0.55	0.09	0.54	0.07
		Inc	0.61	0.14	0.59	0.08	0.54	0.07	0.52	0.06	0.56	0.05	0.57	0.06	0.55	0.07	0.54	0.09
	Angry	Con	0.61	0.14	0.58	0.07	0.54	0.07	0.54	0.07	0.56	0.05	0.57	0.05	0.56	0.07	0.55	0.09
		Inc	0.61	0.12	0.57	0.06	0.53	0.07	0.52	0.05	0.56	0.06	0.56	0.05	0.55	0.09	0.54	0.08
Fast-dm parameters																		
a			0.67	0.10	0.76	0.12	0.70	0.15	0.73	0.11	0.72	0.13	0.81	0.20	0.73	0.10	0.80	0.14
z_r			0.47	0.03	0.49	0.03	0.48	0.02	0.48	0.03	0.48	0.03	0.49	0.04	0.49	0.02	0.48	0.03
v	Happy		0.12	0.26	-0.01	0.19	0.10	0.20	0.06	0.17	0.11	0.27	0.01	0.28	-0.01	0.20	-0.01	0.24
	Sad		0.14	0.22	0.09	0.20	0.01	0.13	0.05	0.15	0.09	0.22	0.03	0.27	0.12	0.22	-0.01	0.24
	Angry		0.18	0.22	0.01	0.25	0.13	0.14	0.09	0.27	0.06	0.24	-0.01	0.20	0.04	0.24	0.14	0.27
t0	Happy		0.44	0.06	0.41	0.06	0.43	0.05	0.40	0.04	0.47	0.09	0.44	0.07	0.44	0.09	0.41	0.06
	Sad		0.44	0.06	0.41	0.07	0.43	0.05	0.40	0.05	0.46	0.09	0.43	0.08	0.44	0.08	0.40	0.07
	Angry		0.44	0.06	0.41	0.07	0.43	0.05	0.41	0.05	0.47	0.09	0.44	0.07	0.45	0.08	0.40	0.07

Table 6

Pre- and post-training descriptive statistics for the dot probe for individuals in the depression cohort who were given the "toward happy" training paradigm

			Active								Sham							
			History n=14				No History n=17				History n=15				No History n=16			
			Pre		Post		Pre		Post		Pre		Post		Pre		Post	
			M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RTs																		
	Happy	Con	0.59	0.08	0.58	0.08	0.54	0.08	0.52	0.05	0.55	0.05	0.56	0.06	0.55	0.09	0.54	0.08
		Inc	0.59	0.08	0.59	0.08	0.54	0.07	0.53	0.06	0.56	0.05	0.56	0.05	0.56	0.09	0.54	0.08
	Sad	Con	0.58	0.09	0.57	0.08	0.54	0.07	0.54	0.06	0.56	0.05	0.55	0.05	0.55	0.09	0.54	0.07
		Inc	0.58	0.08	0.59	0.08	0.54	0.07	0.52	0.06	0.56	0.05	0.57	0.06	0.55	0.07	0.54	0.09
	Angry	Con	0.60	0.09	0.58	0.07	0.54	0.07	0.54	0.07	0.56	0.05	0.57	0.05	0.56	0.07	0.55	0.09
		Inc	0.59	0.10	0.57	0.06	0.53	0.07	0.52	0.05	0.56	0.06	0.56	0.05	0.55	0.09	0.54	0.08
Fast-dm parameters																		
a			0.70	0.11	0.76	0.12	0.70	0.15	0.73	0.11	0.72	0.13	0.81	0.20	0.73	0.10	0.80	0.14
z_r			0.47	0.03	0.49	0.03	0.48	0.02	0.48	0.03	0.48	0.03	0.49	0.04	0.49	0.02	0.48	0.03
v	Happy		0.16	0.26	-0.01	0.19	0.10	0.20	0.06	0.17	0.11	0.27	0.01	0.28	-0.01	0.20	-0.01	0.24
	Sad		0.17	0.19	0.09	0.20	0.01	0.13	0.05	0.15	0.09	0.22	0.03	0.27	0.12	0.22	-0.01	0.24
	Angry		0.13	0.24	0.01	0.25	0.13	0.14	0.09	0.27	0.06	0.24	-0.01	0.20	0.04	0.24	0.14	0.27
t0	Happy		0.43	0.05	0.41	0.06	0.43	0.05	0.40	0.04	0.47	0.09	0.44	0.07	0.44	0.09	0.41	0.06
	Sad		0.43	0.05	0.41	0.07	0.43	0.05	0.40	0.05	0.46	0.09	0.43	0.08	0.44	0.08	0.40	0.07
	Angry		0.43	0.05	0.41	0.07	0.43	0.05	0.41	0.05	0.47	0.09	0.44	0.07	0.45	0.08	0.40	0.07

There were no effects of training on attentional bias for either the anxiety (Table 7), or depression (Table 8) cohorts. The diffusion model parameters did not identify a significant effect of training on decision boundary, starting point, drift rate, or non-decision time in either cohort.

Table 7

Training effects: ANOVA results from the dot probe task in the anxiety cohort

DV	IVs	*df*	*F*	*p*	$\eta^2_{partial}$
RT	TT, TC, Hist, Emo, Trial, Time	1.76, 207.24	0.62	*n.s.*	.01
	TC, Hist, Emo, Trial, Time	1.76, 207.24	0.51	*n.s.*	.004
a	TT, TC, Hist, Time	1, 118	0.57	*n.s.*	.005
	TC, Hist, Time	1, 118	0.37	*n.s.*	.003
z_r	TT, TC, Hist, Time	1, 118	0.02	*n.s.*	<.001
	TC, Hist, Time	1, 118	0.23	*n.s.*	.002
v	TT, TC, Hist, Emo, Time	2, 236	0.14	*n.s.*	.003
	TC, Hist, Emo, Time	2, 236	0.19	*n.s.*	.002
t0	TT, TC, Hist, Emo, Time	2, 236	0.23	*n.s.*	.002
	TC, Hist, Emo, Time	2, 236	2.04	*n.s.*	.01

TT = Training type (toward happy, away angry), TC = Training condition (active, sham), Hist = History (history, no history), Emo = Emotion (happy, sad, angry), Trial = Trial type (congruent, incongruent), Time = Time (pre-training, post training)

* *n.s.* = not significant at the .05 level

Table 8

Training effects: ANOVA results from the dot probe task in the depression cohort

DV	IVs	*df*	*F*	*p*	$\eta^2_{partial}$
RT	TT, TC, Hist, Emo, Trial, Time	2, 236	1.34	*n.s.*	.01
	TC, Hist, Emo, Trial, Time	2, 236	0.41	*n.s.*	.003
a	TT, TC, Hist, Time	1, 118	0.001	*n.s.*	< .001
	TC, Hist, Time	1, 118	0.69	*n.s.*	.01
z_r	TT, TC, Hist, Time	1, 118	1.36	*n.s.*	.01
	TC, Hist, Time	1, 118	0.42	*n.s.*	.004
v	TT, TC, Hist, Emo, Time	2, 236	0.61	*n.s.*	.005
	TC, Hist, Emo, Time	2, 236	1.22	*n.s.*	.01
t0	TT, TC, Hist, Emo, Time	2, 236	1.89	*n.s.*	.02
	TC, Hist, Emo, Time	2, 236	0.99	*n.s.*	.01

TT = Training type (toward happy, away sad), TC = Training condition (active, sham), Hist = History (history, no history), Emo = Emotion (happy, sad, angry), Trial = Trial type (congruent, incongruent), Time = Time (pre-training, post training)

* *n.s.* = not significant at the .05 level

Interpretation bias. Descriptive statistics for the anxiety cohort are presented in Tables 9 and 10. Table 9 reports the mean RTs and fast-dm 30 parameters values and SDs for training away from angry at pre- and post-training; Table 10 reports the mean RTs and fast-dm 30 parameters values and SDs for training toward happy at pre- and post-training. Similarly, descriptive statistics for the depression cohort are presented in Tables 11 and 12. Table 11 reports the mean RTs and fast-dm 30 parameters values and SDs for training away from sad at pre- and post-training; Table 12 reports the mean RTs and fast-dm 30 parameters values and SDs for training toward happy at pre- and post-training.

Table 9

Pre- and post-training descriptive statistics for the yes/no interpretation bias task for individuals in the anxiety cohort who were given the "away from angry" training paradigm

			Active								Sham							
			History n=15				No History n=15				History n=17				No History n=16			
			Pre		Post		Pre		Post		Pre		Post		Pre		Post	
			M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RTs																		
	Happy	Yes	0.86	0.14	0.70	0.16	0.76	0.18	0.68	0.18	0.98	0.45	0.68	0.18	1.02	0.48	0.81	0.33
		No	0.81	0.17	0.73	0.16	0.81	0.17	0.80	0.18	0.87	0.24	0.80	0.35	0.80	0.15	0.74	0.16
	Sad	Yes	1.10	0.39	0.81	0.20	0.99	0.27	0.86	0.30	1.24	0.53	0.89	0.38	1.05	0.43	0.87	0.30
		No	1.03	0.26	0.96	0.25	1.05	0.23	0.89	0.22	1.17	0.58	0.95	0.27	1.14	0.41	0.90	0.22
	Angry	Yes	0.89	0.20	0.73	0.14	0.98	0.34	0.76	0.17	1.06	0.39	0.84	0.35	1.03	0.61	0.85	0.42
		No	1.06	0.32	0.87	0.15	0.91	0.21	0.85	0.21	1.06	0.39	0.91	0.25	1.11	0.41	0.87	0.21
	Neutral	Yes	0.91	0.19	0.83	0.31	0.91	0.23	0.67	0.15	1.05	0.45	0.78	0.37	1.06	0.73	0.88	0.43
		No	0.92	0.19	0.82	0.18	1.04	0.42	0.83	0.17	1.03	0.31	0.79	0.26	1.00	0.34	0.85	0.21
Fast-dm parameters																		
a	Happy		1.77	0.37	1.45	0.39	1.48	0.30	1.41	0.27	1.73	0.53	1.38	0.31	1.71	0.54	1.57	0.52
	Sad		1.70	0.42	1.58	0.34	1.62	0.41	1.34	0.40	1.87	0.58	1.56	0.54	1.75	0.55	1.52	0.56
	Angry		1.78	0.45	1.52	0.36	1.66	0.53	1.48	0.37	1.72	0.48	1.62	0.57	1.84	0.63	1.53	0.48
	Neutral		1.62	0.36	1.56	0.47	1.69	0.46	1.36	0.19	1.72	0.55	1.31	0.50	1.79	0.48	1.49	0.48
zr	Happy		0.63	0.09	0.61	0.10	0.61	0.09	0.64	0.09	0.61	0.09	0.61	0.13	0.61	0.12	0.60	0.09
	Sad		0.57	0.12	0.65	0.12	0.56	0.09	0.58	0.11	0.57	0.13	0.61	0.12	0.61	0.11	0.62	0.07
	Angry		0.60	0.09	0.68	0.09	0.57	0.13	0.64	0.08	0.58	0.13	0.63	0.12	0.57	0.11	0.65	0.09
	Neutral		0.49	0.10	0.58	0.14	0.48	0.15	0.58	0.12	0.54	0.12	0.54	0.12	0.49	0.16	0.53	0.14
v	Happy		-0.11	0.98	0.45	0.71	0.36	0.56	0.32	0.84	0.12	0.80	0.39	1.00	-0.42	1.02	0.18	0.94
	Sad		-0.11	0.60	0.21	0.72	-0.01	0.80	0.19	1.02	-0.21	0.40	0.20	0.89	-0.12	0.65	0.15	0.76
	Angry		0.46	1.02	0.23	0.71	0.43	2.82	0.12	1.09	-0.19	0.71	0.03	0.92	0.13	0.89	-0.21	0.84
	Neutral		-0.14	0.71	-0.22	0.62	0.05	0.67	0.48	0.83	-0.26	1.05	0.01	0.59	0.01	0.60	-0.06	0.91
t0	Happy		0.44	0.09	0.40	0.08	0.47	0.10	0.43	0.13	0.48	0.12	0.44	0.10	0.47	0.09	0.42	0.10

Sad	0.56	0.12	0.47	0.13	0.51	0.12	0.50	0.09	0.52	0.14	0.45	0.12	0.50	0.13	0.46	0.09
Angry	0.46	0.14	0.45	0.10	0.47	0.08	0.45	0.08	0.53	0.11	0.46	0.12	0.46	0.10	0.45	0.04
Neutral	0.47	0.09	0.42	0.16	0.47	0.08	0.40	0.12	0.52	0.20	0.45	0.13	0.46	0.09	0.43	0.08

Table 10

Pre- and post-training descriptive statistics for the yes/no interpretation bias task for individuals in the anxiety cohort who were given the "toward happy" training paradigm

			Active								Sham							
			History n=16				No History n=16				History n=15				No History n=16			
			Pre		Post		Pre		Post		Pre		Post		Pre		Post	
			M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RTs																		
	Happy	Yes	0.81	0.27	0.69	0.15	0.89	0.41	0.65	0.12	0.97	0.24	0.88	0.28	0.76	0.17	0.68	0.17
		No	0.86	0.27	0.78	0.18	0.80	0.23	0.74	0.14	0.82	0.17	0.91	0.41	0.71	0.12	0.72	0.15
	Sad	Yes	1.01	0.27	0.81	0.20	0.95	0.29	0.74	0.15	1.02	0.26	0.92	0.22	0.89	0.26	0.82	0.35
		No	1.01	0.21	0.94	0.26	1.02	0.28	0.86	0.20	1.03	0.21	0.99	0.24	0.89	0.15	0.83	0.18
	Angry	Yes	0.90	0.32	0.76	0.13	0.88	0.15	0.70	0.15	0.99	0.25	0.80	0.10	0.85	0.22	0.74	0.16
		No	1.07	0.32	0.87	0.20	0.90	0.20	0.81	0.25	0.98	0.21	0.95	0.24	0.91	0.26	0.78	0.16
	Neutral	Yes	1.08	0.36	0.79	0.25	0.91	0.29	0.69	0.12	1.10	0.41	0.91	0.30	0.83	0.22	0.69	0.13
		No	1.03	0.35	0.87	0.20	0.94	0.19	0.74	0.07	1.01	0.30	0.94	0.26	0.90	0.25	0.75	0.14
Fast-dm parameters																		
a	Happy		1.60	0.46	1.55	0.35	1.74	0.62	1.33	0.31	1.65	0.34	1.72	0.64	1.42	0.40	1.40	0.32
	Sad		1.81	0.43	1.65	0.56	1.62	0.45	1.44	0.29	1.64	0.34	1.52	0.42	1.46	0.32	1.27	0.28
	Angry		1.78	0.54	1.53	0.39	1.55	0.38	1.44	0.36	1.65	0.44	1.51	0.34	1.51	0.40	1.36	0.49
	Neutral		2.00	0.58	1.44	0.32	1.60	0.46	1.38	0.30	1.69	0.48	1.50	0.36	1.55	0.54	1.35	0.39
zr	Happy		0.60	0.10	0.61	0.09	0.61	0.11	0.58	0.11	0.56	0.11	0.57	0.11	0.60	0.08	0.61	0.11
	Sad		0.59	0.10	0.60	0.11	0.59	0.12	0.63	0.11	0.59	0.12	0.59	0.10	0.58	0.08	0.62	0.11
	Angry		0.59	0.10	0.59	0.09	0.59	0.13	0.58	0.11	0.59	0.12	0.62	0.11	0.64	0.11	0.60	0.11
	Neutral		0.53	0.13	0.56	0.10	0.50	0.11	0.57	0.12	0.48	0.09	0.52	0.09	0.54	0.13	0.59	0.10
v	Happy		0.50	0.85	0.68	0.73	0.14	0.90	0.63	0.97	-0.29	0.73	0.01	0.58	-0.05	0.77	0.26	0.72
	Sad		0.05	0.50	0.27	0.74	0.10	0.67	0.16	1.05	-0.06	0.69	-0.08	0.64	0.05	0.62	-0.03	0.71

	Angry	0.36	0.91	0.42	0.86	-0.23	0.68	0.62	1.11	-0.22	0.56	0.10	0.57	-0.28	0.76	-0.02	0.95
	Neutral	-0.49	0.98	-0.05	0.82	0.23	0.67	-0.07	0.82	-0.21	0.81	-0.21	0.50	0.02	0.87	-0.21	1.00
t0	Happy	0.44	0.09	0.40	0.09	0.43	0.10	0.43	0.07	0.49	0.11	0.42	0.12	0.48	0.07	0.43	0.05
	Sad	0.44	0.09	0.41	0.08	0.50	0.13	0.44	0.10	0.53	0.10	0.49	0.11	0.48	0.09	0.50	0.09
	Angry	0.41	0.08	0.43	0.07	0.50	0.08	0.41	0.11	0.51	0.12	0.48	0.09	0.48	0.08	0.46	0.07
	Neutral	0.45	0.15	0.42	0.07	0.47	0.15	0.39	0.06	0.48	0.07	0.46	0.11	0.44	0.11	0.44	0.08

Table 11

Pre- and post-training descriptive statistics for the yes/no interpretation bias task for individuals in the depression cohort who were given the "away from sad" training paradigm

			Active								Sham							
			History n=16				No History n=17				History n=14				No History n=16			
			Pre		Post		Pre		Post		Pre		Post		Pre		Post	
			M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RTs																		
	Happy	Yes	0.91	0.38	0.68	0.20	0.75	0.18	0.72	0.18	0.84	0.27	0.71	0.13	0.77	0.18	0.68	0.15
		No	0.84	0.19	0.72	0.14	0.73	0.14	0.73	0.13	0.84	0.18	0.72	0.17	0.77	0.15	0.73	0.17
	Sad	Yes	1.03	0.31	0.71	0.22	0.94	0.21	0.83	0.14	1.06	0.26	0.82	0.25	0.89	0.20	0.75	0.11
		No	1.20	0.58	0.89	0.28	0.99	0.29	0.95	0.23	1.07	0.26	0.97	0.33	0.87	0.16	0.81	0.17
	Angry	Yes	1.05	0.30	0.71	0.18	0.83	0.15	0.79	0.18	0.98	0.19	0.74	0.20	0.85	0.13	0.73	0.12
		No	0.96	0.23	0.87	0.26	0.89	0.18	0.86	0.17	0.90	0.19	0.88	0.22	0.90	0.18	0.79	0.15
	Neutral	Yes	0.91	0.25	0.76	0.25	0.87	0.28	0.78	0.21	1.01	0.39	0.78	0.20	0.82	0.20	0.69	0.20
		No	1.05	0.33	0.86	0.18	0.92	0.23	0.82	0.12	0.95	0.31	0.85	0.18	0.87	0.17	0.78	0.13
Fast-dm parameters																		
a	Happy		1.78	0.53	1.41	0.41	1.59	0.46	1.43	0.41	1.59	0.45	1.47	0.28	1.52	0.33	1.33	0.26
	Sad		1.83	0.68	1.32	0.40	1.58	0.44	1.46	0.38	1.73	0.32	1.61	0.44	1.45	0.36	1.35	0.36
	Angry		1.74	0.52	1.46	0.43	1.51	0.43	1.50	0.46	1.61	0.39	1.53	0.43	1.54	0.33	1.38	0.38
	Neutral		1.66	0.57	1.54	0.54	1.65	0.58	1.44	0.36	1.72	0.44	1.48	0.29	1.42	0.29	1.26	0.27
zr	Happy		0.62	0.12	0.55	0.11	0.61	0.12	0.63	0.10	0.54	0.14	0.57	0.15	0.63	0.12	0.61	0.10
	Sad		0.57	0.09	0.63	0.09	0.57	0.14	0.57	0.07	0.59	0.11	0.60	0.13	0.55	0.10	0.54	0.15
	Angry		0.61	0.12	0.62	0.11	0.62	0.06	0.64	0.11	0.55	0.11	0.59	0.15	0.62	0.12	0.62	0.11

	Neutral	0.49	0.13	0.53	0.14	0.51	0.14	0.53	0.11	0.45	0.11	0.56	0.12	0.53	0.13	0.51	0.12
v	Happy	0.07	0.72	0.65	0.61	0.22	0.81	0.23	1.03	0.20	0.82	0.31	1.16	-0.11	0.70	0.32	1.02
	Sad	0.15	0.55	-0.09	0.82	0.09	0.91	0.07	0.83	0.00	0.64	0.40	0.93	-0.01	0.57	0.40	0.66
	Angry	-0.28	0.69	0.18	0.69	-0.14	0.58	0.00	0.80	-0.15	0.73	0.56	0.87	-0.24	0.52	0.24	0.85
	Neutral	0.22	0.84	0.24	1.00	0.12	0.64	0.15	0.58	0.07	0.81	-0.17	0.77	-0.17	0.87	0.30	0.97
t0	Happy	0.44	0.08	0.40	0.09	0.44	0.07	0.45	0.07	0.48	0.13	0.42	0.12	0.46	0.07	0.42	0.05
	Sad	0.45	0.12	0.44	0.13	0.52	0.07	0.48	0.06	0.53	0.09	0.42	0.15	0.47	0.11	0.45	0.06
	Angry	0.48	0.10	0.41	0.12	0.46	0.06	0.45	0.08	0.53	0.12	0.41	0.10	0.47	0.06	0.44	0.08
	Neutral	0.47	0.12	0.38	0.13	0.44	0.11	0.42	0.08	0.48	0.13	0.45	0.12	0.48	0.10	0.41	0.08

Table 12

Pre- and post-training descriptive statistics for the yes/no interpretation bias task for individuals in the depression cohort who were given the "toward happy" training paradigm

			Active								Sham							
			History n=14				No History n=17				History n=17				No History n=15			
			Pre		Post		Pre		Post		Pre		Post		Pre		Post	
			M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
RTs																		
	Happy	Yes	0.81	0.30	0.70	0.12	0.85	0.40	0.65	0.18	0.92	0.34	0.71	0.17	0.71	0.14	0.66	0.16
		No	0.86	0.14	0.84	0.11	0.73	0.14	0.74	0.20	0.89	0.35	0.75	0.15	0.72	0.11	0.87	0.41
	Sad	Yes	0.91	0.26	0.81	0.25	0.91	0.35	0.78	0.23	1.21	0.45	0.97	0.26	0.90	0.19	0.80	0.17
		No	1.03	0.21	0.94	0.13	0.99	0.30	0.86	0.19	1.19	0.68	1.02	0.35	0.92	0.12	0.87	0.18
	Angry	Yes	0.89	0.19	0.80	0.13	0.87	0.22	0.70	0.18	1.06	0.31	0.86	0.24	0.87	0.21	0.74	0.14
		No	0.95	0.21	0.91	0.13	0.91	0.21	0.82	0.20	1.06	0.24	0.99	0.39	0.91	0.17	0.82	0.13
	Neutral	Yes	0.92	0.31	0.78	0.25	0.78	0.20	0.65	0.14	1.07	0.35	0.83	0.21	0.85	0.16	0.71	0.14
		No	1.00	0.24	0.89	0.14	0.96	0.32	0.83	0.17	1.05	0.33	0.92	0.24	0.87	0.18	0.81	0.21
Fast-dm parameters																		
a	Happy		1.82	0.52	1.67	0.16	1.58	0.53	1.51	0.31	1.74	0.53	1.50	0.40	1.50	0.40	1.46	0.45
	Sad		1.70	0.38	1.59	0.37	1.56	0.46	1.43	0.31	1.75	0.57	1.77	0.95	1.60	0.35	1.39	0.26
	Angry		1.66	0.40	1.62	0.44	1.55	0.43	1.40	0.37	1.88	0.47	1.70	0.48	1.55	0.34	1.42	0.21
	Neutral		1.72	0.47	1.52	0.39	1.49	0.39	1.37	0.33	1.91	0.67	1.64	0.42	1.49	0.34	1.33	0.34
zr	Happy		0.65	0.11	0.64	0.15	0.60	0.14	0.66	0.13	0.58	0.09	0.55	0.12	0.60	0.10	0.57	0.11

	Sad	0.63	0.13	0.66	0.09	0.59	0.13	0.60	0.15	0.56	0.09	0.56	0.15	0.57	0.09	0.57	0.10
	Angry	0.65	0.06	0.67	0.11	0.62	0.11	0.67	0.06	0.57	0.10	0.55	0.12	0.62	0.12	0.67	0.09
	Neutral	0.51	0.10	0.55	0.18	0.53	0.07	0.56	0.12	0.47	0.15	0.49	0.09	0.51	0.13	0.56	0.13
v	Happy	0.48	1.14	0.56	0.84	-0.01	0.92	0.16	1.10	0.23	0.65	0.95	0.96	0.24	0.94	0.82	0.96
	Sad	0.21	1.30	0.49	1.53	0.18	0.53	0.11	0.85	-0.15	0.57	-0.01	0.81	0.11	0.80	0.09	0.74
	Angry	-0.21	0.53	-0.13	0.94	-0.24	0.75	0.11	1.03	0.02	0.72	0.52	1.19	-0.10	0.83	-0.20	0.76
	Neutral	0.05	0.74	0.34	1.28	0.39	0.53	0.37	0.68	0.07	0.84	0.24	0.55	-0.17	0.77	-0.04	1.10
t0	Happy	0.45	0.06	0.42	0.10	0.47	0.10	0.42	0.09	0.44	0.11	0.42	0.10	0.43	0.08	0.39	0.10
	Sad	0.50	0.08	0.46	0.09	0.46	0.12	0.48	0.08	0.52	0.12	0.48	0.08	0.44	0.10	0.43	0.14
	Angry	0.49	0.09	0.48	0.09	0.50	0.10	0.44	0.08	0.47	0.12	0.41	0.10	0.48	0.11	0.45	0.07
	Neutral	0.45	0.12	0.44	0.09	0.44	0.07	0.41	0.07	0.48	0.11	0.40	0.12	0.46	0.08	0.41	0.12

Analysis of RTs in the anxiety cohort identified an interaction between training, history, emotion, response and time (Table 13). However, follow-up analyses did not identify a training effect on yes responses that indicated a shift in bias. An interaction between training type, condition, history, emotion and time was also identified for boundary separation. However, as with RT analysis, follow-up analyses did not identify a training effect on boundary separation. There was no effect of training on starting point, drift rate, or non-decision time.

Table 13

Training effects: ANOVA results from the yes/no interpretation bias task in the anxiety cohort

DV	IVs	*df*	*F*	*p*	$\eta^2_{partial}$
RT	TT, TC, Hist, Emo, Resp, Time	3, 354	0.24	*n.s.*	<.001
	TC, Hist, Emo, Trial, Time	3, 354	3.16	= .03	.03
a	TT, TC, Hist, Emo, Time	3, 354	2.85	= .04	.02
	TC, Hist, Emo, Time	3, 354	1.19	*n.s.*	.01
z_r	TT, TC, Hist, Emo, Time	3, 354	0.66	*n.s.*	.001
	TC, Hist, Emo, Time	3, 354	0.18	*n.s.*	.001
v	TT, TC, Hist, Emo, Time	2.70, 318.04	2.31	*n.s.*	.02
	TC, Hist, Emo, Time	2.70, 318.04	1.36	*n.s.*	.01
t0	TT, TC, Hist, Emo, Time	3, 354	1.06	*n.s.*	.01

	TC, Hist, Emo, Time	3, 354	1.70	*n.s.*	.01

TT = Training type (toward happy, away angry), TC = Training condition (active, sham), Hist = History (history, no history), Emo = Emotion (happy, sad, angry, neutral), Trial = Trial type (yes, no), Time = Time (pre-training, post training)
* *n.s.* = not significant at the .05 level

For the depression cohort, there were no effects of training on interpretation bias. (Table 14). Additionally, the diffusion model parameters did not identify a significant effect of training on decision boundary, starting point, or drift rate. While there was a significant interaction detected for raining type, condition, history, emotion and time on non-decision time, follow-up analyses failed to identify any effect of training on non-decision time.

Table 14

Training effects: ANOVA results from the yes/no interpretation bias task in the depression cohort

DV	IVs	*df*	*F*	*p*	$\eta^2_{partial}$
RT	TT, TC, Hist, Emo, Trial, Time	3, 354	0.68	*n.s.*	.01
	TC, Hist, Emo, Trial, Time	3, 354	0.39	*n.s.*	< .001
a	TT, TC, Hist, Emo, Time	2.69, 317.58	0.93	*n.s.*	.01
	TC, Hist, Emo, Time	2.69, 317.58	1.04	*n.s.*	.01
z_r	TT, TC, Hist, Emo, Time	3, 354	0.35	*n.s.*	.003

		TC, Hist, Emo, Time	3, 354	1.22	*n.s.*	.01
v		TT, TC, Hist, Emo,	2.80,	1.14	*n.s.*	.01
	Time		330.76			
		TC, Hist, Emo, Time	2.80,	1.46	*n.s.*	.01
			330.76			
t0		TT, TC, Hist, Emo,	3, 354	3.28	= .02	.03
	Time					
		TC, Hist, Emo, Time	3, 354	1.46	*n.s.*	.01

TT = Training type (toward happy, away sad), TC = Training condition (active, sham), Hist = History (history, no history), Emo = Emotion (happy, sad, angry, neutral), Trial = Trial type (yes, no), Time = Time (pre-training, post training)

**n.s.* = not significant at the .05 level

Reliability Analyses

To examine the test-retest reliability of the traditional attentional bias scores, and diffusion model parameters from the dot probe task, we conducted correlational analyses between the values obtained at pre-training and post training for individuals in the sham training condition. As can be seen in table 15, correlations for the attentional bias score were not positively correlated from pre- to post-training, revealing no evidence of test-retest reliability. For the diffusion model parameters, non-decision time values for pre- and post-training were highly correlated for all conditions. The boundary separation values for pre- and post-training were moderately correlated. Starting point and drift rate values for pre- and post-training were not correlated.

Table 15

Test-retest correlation coefficients obtained from the dot probe task for the sham training groups

	Anxiety	Depression

Attentional Bias		
Happy	-.1	-.15
Sad	-.44**	-.09
Angry	-.13	-.11
zr	.12	.18
a	.67**	.49**
v Happy	-.04	.07
v Angry	.06	.04
v Sad	.11	-.01
t0 Happy	.83**	.82**
t0 Sad	.75**	.82**
t0 Angry	.80**	.80**

**Correlation is significant at the 0.01 level (2-tailed).

*Correlation is significant at the 0.05 level (2-tailed).

Discussion

The aim of this study was two-fold. The first aim was to try to identify the cognitive processes that underlie negative attentional and interpretation biases in individuals with anxiety and depression. The second aim was to identify which cognitive processes shift after successful attentional bias modification. Based on previous attentional bias research, we expected that individuals with a history of anxiety would have an attentional bias for angry faces, while individuals with a history of depression would have an attentional bias for sad faces. We expected that these would be evident from attentional bias scores, as well as the diffusion model parameter of drift rate. We also hypothesised that there may be associated non-decisional differences due to hypervigilance and motor slowing, which are symptoms of anxiety and depression, respectively

(DSM-V; American Psychiatric Association, 2014). Further, we expected that interpretation bias to disorder congruent stimuli would be evident from faster response times to "yes" responses, and from higher drift rates and starting point values for angry and sad faces for individuals with a history of anxiety and depression, respectively.

The Absence of Attentional Bias

The results derived from the dot probe assessment task and the yes/no task did not identify disorder-congruent attentional or interpretation biases in participants with a history of anxiety or depression. Therefore, our expectation that we would identify disorder-congruent attentional and interpretation biases in individuals with a history of anxiety and depression was not supported. Further, to identify the cognitive processes of change that occur because of successful attentional bias modification, an attentional bias must be captured in the first instance. In line with the gold standard of intervention studies, the condition that is under investigation must be present prior to assessing success of treatment. Given the premise is that individuals with anxiety and/or depression possess these cognitive biases, not seeing them in the first instance makes any changes thereafter irrelevant therapeutically. As a result, our goal to identify the cognitive processes of change was unable to be achieved. There are several possibilities that we need to consider as to why we did not find an attentional bias in our sample. The first is sample characteristics. The second is stimulus properties. The third is that the dot probe task may be an unreliable measure of attentional bias. The fourth is that attentional bias may not be a stable trait. Finally, attentional bias may not be able to be effectively, or consistently, measured with the dot-probe task in clinical populations. We address each of these explanations in turn.

A non-clinical sample was recruited for this study. To classify participants into the respective clinical groups, we asked participants if they had a history of anxiety or depression; however, we did not clarify how long ago they had experienced the disorder, or its severity. Therefore, we need to consider whether our sample did not have enough clinical severity to identify an attentional bias. The pre-training DASS scores demonstrate that participants who reported a

history of anxiety or depression had higher levels of anxiety, depression and emotion regulation difficulties than participants who reported no such history. However, the mean scores on the anxiety and depression measures for those with a history of anxiety and depression were in only in the moderate to high range. Scores above 12 on the stress subscale and scores above 10 on the depression subscale would be needed to classify symptoms in the severe range. There is some evidence that attentional bias is difficult to capture, or does not exist, in mild to moderate cases (e.g. Broadbent & Broadbent, 1988). To check whether this could account for the lack of bias scores found in our sample, we examined attentional bias in participants who had clinical scores in the severe range. The subsample of 35 participants who had an anxiety (stress subscale) score of 13 or above showed no attentional bias for angry stimuli, nor did the bias for this subsample differ from that of those who had no or low levels of anxiety. Similarly, the 39 participants who had a depression score of 11 and above, showed no attentional bias for sad stimuli, nor did the bias for this subsample differ from that of those who had no or low levels of depression. Considering the limitations of lower power associated with smaller sample sizes, the lack of an attentional bias identified in these subsamples indicates that the reason we did not find an attentional bias does not appear to be due to low clinical scores resulting from recruiting a non-clinical sample.

Additionally, effects of attentional bias are generally small to moderate (eg. Bar Haim, 2010; Browning et al, 2010), as are effects of attentional bias modification. As there was no attentional bias captured in the current study, nor was there an effect of training detected, it may be that a larger sample size is required to detect small to moderate effects in clinical populations than was presented here.

Turning to the stimuli we used, physical properties, such as colour, shape, and configuration, of the stimuli chosen can affect attentional responding (Turatto & Galfano, 2000). For example, stimulus location (top/bottom, left/right), and stimulus-type (pictures, words, abstract) can moderate the strength of attentional bias (Beard et al., 2012). We used faces from the NimStim database, which have successfully captured attention in other studies (e.g. Arndt & Fujiwara, 2012; Niles,

Mesri, Burklund, Lieberman, & Craske, 2013; Paulewicz, Blaut, & Kłosowska, 2012). Therefore, stimulus properties cannot explain the lack of an attentional bias in our sample.

The attentional bias score derived from the dot probe task has poor reliability (Cisler et al., 2009; Rodebaugh et al., 2016; Schmukle, 2005), and could account for the lack of an attentional bias in this sample based on RT measures. However, the lack of reliability of the score cannot account for not capturing attentional bias with the diffusion model parameters. Assuming the bias is there, and the unreliability of the attentional bias score is the problem, as the diffusion model parameters separate out the cognitive processes captured by the RT score, reducing noise in the data, theoretically, it should offer a more precise account of attentional bias. Thus, if an attentional bias was there but simply not captured due to a lack of reliability of the attentional bias score the diffusion model should be able to identify it. As it did not, we turn to the consideration that attentional bias may not be a stable trait.

The variability in the successful capture of attentional bias in clinical populations may indicate that attentional bias is not a stable trait, but rather a dynamic process that fluctuates over time (Rodebaugh et al., 2016). This theory of a dynamic attentional bias proposes that attentional bias presents in phasic bursts, oriented towards and then away from salient stimuli (Zvielli, Bernstein, & Koster, 2015). Further, this dynamic process has been posited to account for the lack of reliability often found in the attentional bias score (Rodebaugh et al., 2016). That is, an attentional bias may be captured at one time point, but not another, depending on the dynamic processes that are activated at the time of assessment, and thus the two attentional bias scores do not correlate. This explanation may account for the lack of an attentional bias in this sample.

If attentional bias is a stable trait with a causal or maintaining role in anxiety and depression, we would have found it in the subset of particiapnts with high levels of anxiety and depression. Additionally, if attentional bias was a stable trait and it was the reliability of the attentional bias score that was the issue, we would still have captured attentional bias through the diffusion model parameters. Given that neither the diffusion model parameters or the attentional bias score captured

an attentional bias, it appears that a dynamic temporal of attentional bias is the most plausible explanation for why we did not find an attentional bias in our sample.

However, there is a more sobering explanation that must be considered. It may be that the dot-probe task is unable to accurately or reliably detect attentional bias effects in disorders with such broad manifestations as anxiety and depression. If the effect is not able to be detected in the behavioural data captured by the task, it will not be identified by any form of computational model, no matter how stringent.

Differences in Attentional and Interpretation Processes Between Anxiety and Depression

The lack of an attentional bias in our sample has voided our primary aim of identifying the cognitive processes responsible for attentional and interpretation bias. Therefore, our attention now turns to the findings of the cognitive processes that differ between anxiety and depression. Analysis of fast-dm data demonstrated activation of slightly different cognitive processes for these two disorders. Importantly, these processes were not detected by RTs, demonstrating the sensitivity of the diffusion model parameters. We first examine how attentional processes differ between anxiety and depression, as well as the added information we obtained from fast-dm 30 analysis relative to RT analysis.

RTs from the dot probe task did not identify any difference between congruent responses and incongruent responses for any emotion, irrespective of clinical history. This finding would lead us to conclude that there was no capture by emotional stimuli in our sample. However, the diffusion model analysis identified that in the groups with a history of anxiety and a history of depression, that evidence for the probe was stronger when presented behind an emotive stimulus than when presented behind a neutral stimulus, indicating that attention *was* captured by emotion. Additionally, individuals with a history of depression had slower RTs than individuals with no history of depression, which we may have hypothesised to be a function of motor slowing, a common symptom of depression. However, fast-dm 30 analysis identified that individuals with a history of depression had faster non-decisional processes than individuals with no history of

depression. This finding informs us that the slower RTs for individuals with a history of depression is not because of motor slowing, and that they in fact appear to have a phasic response to emotional stimuli. Therefore, diffusion model analysis identified faster emotional capture of attention in individuals with a history of anxiety and depression than individuals with no such history. Additionally, individuals with depression appear to have a phasic response to emotion.

Turning our focus to the processes identified in the interpretation of stimuli, response times from the yes/no task did not identify differences in the interpretation of emotional stimuli based on clinical history. RTs did identify a tendency across all participants to respond more quickly to happy stimuli. The parameters returned by fast-dm 30 identified that these faster RTs were a result of less conservative responding, and faster non-decisional processes. Individuals with a history of depression had slower mean RTs than individuals with no history of depression, for which diffusion model parameters identified to be a result of more conservative decision-making than individuals with no history of depression; an alternative account for their longer RTs than the motor slowing hypothesis. Additionally, and unexpectedly, individuals with a history of depression had less expectancy for angry information than individuals with no history of depression. While the function of this result is not clear, one potential explanation for this is that depression is often associated with an avoidant coping style. For individuals with depression, anger can be associated with rejection, rather than threat (Leyman, De Raedt, Schacht, & Koster, 2007), and as such, the lower starting point may be indicative of efforts to avoid engaging with angry information.

The decisional processes for interpretation of emotion for individuals with a history of anxiety were slightly different. While RTs identified no differences between individuals with a reported history of anxiety and those with no such history, the parameters returned from fast-dm 30 analysis identified that individuals with anxiety had a propensity to accumulate evidence more quickly across all emotions than individuals without a history of anxiety. This could be a function of hypervigilance. Individuals with anxiety may analyse and interpret the stimuli in their environment more quickly than individuals with no such history due to a heightened fear of threat.

Therefore, there were two primary differences in the decision-making processes associated with interpreting affective material for individuals with a history of anxiety and depression. Individuals with depression were slower to respond due to more cautious responding, i.e., they need more evidence before they will execute a decision, whereas individuals with a history of anxiety were faster to gather evidence to make their decision. While we were unable to identify any effects of training on attentional and interpretation bias in this study, we did demonstrate that the diffusion model can identify differences in cognitive processing between individuals with anxiety and depression in a way that RTs cannot.

Reliability of the Diffusion Model Parameters for the Attentional Bias Task

Our second goal of this study was to assess the test-retest reliability of the attentional bias parameters of the diffusion model returned by Fast-dm. We wanted to assess whether the diffusion model parameters returned for the dot probe task could be a more stable and reliable measure of attentional bias than the dot probe score. To do so, we used test/re-test reliability, which may be open to practise effects. This is a potential limitation of this method of reliability assessment. We found that neither the attentional bias score, starting point, or drift rate values correlated from pre-training to post training. However, values from pre- to post training were highly correlated for non-decision time, and moderately correlated for boundary separation.

We have hypothesised that attentional bias is likely a processing bias, which would be captured by the drift rate parameter, therefore, high test-retest reliability of the drift rate value is paramount. However, before we dismiss the diffusion model analysis as being unsuitable based on the correlations presented above, we consider how the lack of variability of drift rate values in our sample may account for the lack of correlation for these values at pre- and post-training. A lack of variability in data points can reduce the likelihood of a correlation, or decrease its strength (Goodwin & Leech, 2006). Drift rate values can typically range from -5 to +5, which is a range of 10. In our sample, the range for drift rate values was just above 1. This small range is primarily because we did not capture an attentional bias in this sample. Whether this is due to poor quality

stimuli, the potential dynamic variability of attentional bias (both discussed above), or because attentional bias is not a function of perceptual bias, the low values, and small variance obtained from the drift rate parameter, have decreased the likelihood of a meaningful correlation. Therefore, to determine whether the drift rate is reliable across time, analysis of data that did identify an attentional bias would be needed to determine which of these factors is responsible for the lack of correlation found in this study.

Theoretical and Clinical Implications

The findings presented in this study have several important theoretical implications. The first is that our results support Zvielli, Bernstein, and Koster's (2015) notion that attentional bias may not be a stable trait, but instead is dynamic in nature. Despite the high levels of anxiety and depression reported by some participants, we found no evidence of an attentional bias through RT analysis.

Second, our results illustrate the potential for the diffusion model to identify processes in attentional and interpretation bias that are not captured by RTs alone. Specifically, we identified that the processes involved in attention and interpretation differ slightly for individuals with a history of anxiety and depression. Further, these processes differ from those of individuals with no such history. The identification of differences in decisional processes for attention and interpretation means that if the diffusion model was used to analyse data that has captured an attentional bias, the decisional processes responsible for that bias could be identified.

The clinical implications of these findings are promising. Enhanced understanding of the decisional processes that are responsible for affective cognitive biases offers the opportunity to develop and implement more targeted strategies for cognitive change. This, in turn, has the potential to alleviate the emotional distress that often accompanies these distorted cognitive processes.

Conclusion

Our goal of identifying the processes involved in successful attentional bias modification, and the impact ABM may have on interpretation bias for individuals with anxiety and depression,

was not met. We did not identify any disorder-congruent biases. However, our analyses did identify that the parameters of fast-dm 30 have the capacity to identify cognitive processes that are masked by RTs alone. Further, differences were identified in cognitive processes between clinical cohorts. The implications of this for enhancing our theoretical understanding of the decisional processes that are hypothesised to contribute to the onset and maintenance of anxiety and depression is promising. However, the test-retest reliability of the parameters for both the dot-probe task and the yes/no task were mixed. Further research needs to ensure that the methodological issues that can impact on reliability are accounted for before we can make firm conclusions as to the diffusion model's viability as an alternative, more reliable, way to analyse data from the dot probe task.

References

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders-DSM-5. *Washington, DC: Author*.

Amir, N., Beard, C., Burns, M., & Bomyea, J. (2009). Attention modification program in individuals with generalized anxiety disorder. *Journal of Abnormal Psychology, 118*(1), 28.

Arndt, J. E., & Fujiwara, E. (2012). Attentional bias towards angry faces in trait-reappraisal. *Personality and Individual Differences*, *52*(1), 61–66. https://doi.org/10.1016/j.paid.2011.08.030

Bar-Haim, Y. (2010). Research review: attention bias modification (ABM): a novel treatment for anxiety disorders. *Journal of Child Psychology and Psychiatry*, *51*(8), 859–870.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. International Universities Press.

Broadbent, D., & Broadbent, M. (1988). Anxiety and attentional bias: State and trait. *Cognition & Emotion*, *2*(3), 165–183.

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: results from a pilot study. *J Behav Ther Exp Psychiatry*, *42*, 258–64. https://doi.org/10.1016/j.jbtep.2010.12.006

Browning, M., Holmes, E. A., & Harmer, C. J. (2010). The modification of attentional bias to emotional information: A review of the techniques, mechanisms, and relevance to emotional disorders. *Cogn Affect Behav Neurosci*, *10*, 8–20. https://doi.org/10.3758/CABN.10.1.8

Burcusa, S. L., & Iacono, W. G. (2007). Risk for Recurrence in Depression. *Clinical Psychology Review*, *27*(8), 959–985. https://doi.org/10.1016/j.cpr.2007.02.005

Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016). Scaling-up treatment of depression and anxiety: a global return on investment analysis. *The Lancet. Psychiatry*, *3*(5), 415–424. https://doi.org/10.1016/S2215-0366(16)30024-4

Cisler, J. M., & Olatunji, B. O. (2012). Emotion Regulation and Anxiety Disorders. *Current Psychiatry Reports*, *14*(3), 182–187. https://doi.org/10.1007/s11920-012-0262-2

Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & Emotion*, *6*(6), 409–434.

Eysenck, M. W., 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*(2), 144.

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General*, *136*(3), 389.

Goodwin, L. D., & Leech, N. L. (2006). Understanding Correlation: Factors That Affect the Size of r. *The Journal of Experimental Education*, *74*(3), 249–266. https://doi.org/10.3200/JEXE.74.3.249-266

Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: current status and future directions. *Annual Review of Clinical Psychology*, *6*, 285.

Gratz, K. L., & Roemer, L. (2004a). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment*, *26*(1), 41–54.

Gratz, K. L., & Roemer, L. (2004b). Multidimensional Assessment of Emotion Regulation and Dysregulation: Development, Factor Structure, and Initial Validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment*, *26*(1), 41–54. https://doi.org/10.1023/B:JOBA.0000007455.08539.94

Gross, J. J. (2013). Emotion regulation: Taking stock and moving forward. *Emotion*, *13*, 359–365. https://doi.org/10.1037/a0032135

Gross, J. J., & Thompson, R. A. (2007). Emotion regulation: Conceptual foundations. Retrieved from http://psycnet.apa.org/psycinfo/2007-01392-001

Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychol Bull*, *137*, 940–58. https://doi.org/10.1037/a0024355

Heathcote, A., Suraev, A., Curley, S., Gong, Q., Love, J., & Michie, P. T. (2015). *Decision processes and the slowing of simple choices in schizophrenia.* American Psychological Association. Retrieved from http://psycnet.apa.org/journals/abn/124/4/961/

Joormann, J., & Quinn, M. E. (2014). Cognitive processes and emotion regulation in depression. *Depress Anxiety*, *31*(4), 308–15. https://doi.org/10.1002/da.22264

Joormann, Jutta, & Siemer, M. (2011). Affective processing and emotion regulation in dysphoria and depression: Cognitive biases and deficits in cognitive control. *Social and Personality Psychology Compass*, *5*(1), 13–28.

Lerche, V., & Voss, A. (2016). Model Complexity in Diffusion Modeling: Benefits of Making the Model More Parsimonious. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01324

Leyman, L., De Raedt, R., Schacht, R., & Koster, E. H. W. (2007). Attentional biases for angry faces in unipolar depression. *Psychological Medicine*, *37*(3), 393–402. https://doi.org/10.1017/S003329170600910X

Lovibond, S. H., & Lovibond, P. F. (1996). *Manual for the depression anxiety stress scales*. Sydney : Psychology Foundation of Australia. Retrieved from http://trove.nla.gov.au/version/38629420

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, 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. https://doi.org/10.1037//0021-843x.111.1.107

Mogg, K., & Bradley, B. P. (1999). Some methodological issues in assessing attentional biases for threatening faces in anxiety: A replication study using a modified version of the probe detection task. *Behaviour Research and Therapy*, *37*(6), 595–604.

Mogg, K., & Bradley, B. P. (2016). Anxiety and attention to threat: Cognitive mechanisms and treatment with attention bias modification. *Behaviour Research and Therapy*, *87*(Supplement C), 76–108. https://doi.org/10.1016/j.brat.2016.08.001

Niles, A. N., Mesri, B., Burklund, L. J., Lieberman, M. D., & Craske, M. G. (2013). Attentional bias and emotional reactivity as predictors and moderators of behavioral treatment for social phobia. *Behaviour Research and Therapy*, *51*(10), 669–679. https://doi.org/http://dx.doi.org/10.1016/j.brat.2013.06.005

Nosek, B. A., & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition*, *19*(6), 625–666. https://doi.org/10.1521/soco.19.6.625.20886

O'Leary, M. J., & Barber, P. J. (1993). Interference effects in the Stroop and Simon paradigms. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(4), 830.

Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*, *5*(5), 376–404. https://doi.org/10.1167/5.5.1

Paulewicz, B., Blaut, A., & Kłosowska, J. (2012). Cognitive effects of attentional training depend on attentional control. *Polish Psychological Bulletin*, *43*(4), 272–277. https://doi.org/10.2478/v10059-012-0030-5

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.

Renaud, J., Russell, J. J., & Myhr, G. (2014). Predicting Who Benefits Most From Cognitive-Behavioral Therapy for Anxiety and Depression. *Journal of Clinical Psychology*, *70*(10), 924–932. https://doi.org/10.1002/jclp.22099

Rodebaugh, T. L., Scullin, R. B., Langer, J. K., Dixon, D. J., Huppert, J. D., Bernstein, A., … Lenze, E. J. (2016). Unreliability as a threat to understanding psychopathology: The

cautionary tale of attentional bias. *Journal of Abnormal Psychology, 125*(6), 840–851. https://doi.org/10.1037/abn0000184

Schoenmakers, T., Wiers, R. W., Jones, B. T., Bruce, G., & Jansen, A. T. M. (2007). Attentional re-training decreases attentional bias in heavy drinkers without generalization. *Addiction (Abingdon, England), 102*(3), 399–405. https://doi.org/10.1111/j.1360-0443.2006.01718.x

Scholten, W. D., Batelaan, N. M., van Balkom, A. J., WJH. Penninx, B., Smit, J. H., & van Oppen, P. (2013). Recurrence of anxiety disorders and its predictors. *Journal of Affective Disorders, 147*(1), 180–185. https://doi.org/10.1016/j.jad.2012.10.031

Schulz, K. P., Fan, J., Magidina, O., Marks, D. J., Hahn, B., & Halperin, J. M. (2007). Does the emotional go/no-go task really measure behavioral inhibition?: Convergence with measures on a non-emotional analog. *Archives of Clinical Neuropsychology, 22*(2), 151–160. https://doi.org/10.1016/j.acn.2006.12.001

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. *J Abnorm Psychol, 118*, 65–75. https://doi.org/10.1037/a0014377

Shechner, T., Britton, J. C., Pérez-Edgar, K., Bar-Haim, Y., Ernst, M., Fox, N. A., … Pine, D. S. (2012). Attention biases, anxiety, and development: toward or away from threats or rewards? *Depression and Anxiety, 29*(4), 282–294. https://doi.org/10.1002/da.20914

Taylor, C. T., Bomyea, J., & Amir, N. (2011). Malleability of attentional bias for positive emotional information and anxiety vulnerability. *Emotion, 11*(1), 127.

Tottenham, N., Hare, T. A., & Casey, B. J. (2011). Behavioral assessment of emotion discrimination, emotion regulation, and cognitive control in childhood, adolescence, and adulthood. *Front Psychol, 2*, 39. https://doi.org/10.3389/fpsyg.2011.00039

Tottenham, Nim, Hare, T. A., & Casey, B. (2011). Behavioral assessment of emotion discrimination, emotion regulation, and cognitive control in childhood, adolescence, and adulthood. *Frontiers in Psychology, 2*.

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods, 39*(4), 767–775. https://doi.org/10.3758/BF03192967

Voss, A., Voss, J., & Lerche, V. (2015a). Assessing cognitive processes with diffusion model analyses: a tutorial based on fast-dm-30. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00336

Voss, A., Voss, J., & Lerche, V. (2015b). Assessing cognitive processes with diffusion model analyses: a tutorial based on fast-dm-30. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00336

Waters, A. M., Pittaway, M., Mogg, K., Bradley, B. P., & Pine, D. S. (2013). Attention training towards positive stimuli in clinically anxious children. *Developmental Cognitive Neuroscience, 4*, 77–84. https://doi.org/10.1016/j.dcn.2012.09.004

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Using diffusion models to understand clinical disorders. *J Math Psychol, 54*, 39–52. https://doi.org/10.1016/j.jmp.2010.01.004

White, Corey N., Skokin, K., Carlos, B., & Weaver, A. (2016). Using decision models to decompose anxiety-related bias in threat classification. *Emotion, 16*(2), 196.

Zvielli, A., Bernstein, A., & Koster, E. H. (2015). Temporal dynamics of attentional bias. *Clinical Psychological Science, 3*(5), 772–788.

Chapter 6: Conclusion

The aims of this thesis were threefold. The first aim was to successfully fit the diffusion model to data from the dot probe task. The second was to contribute to the broader understanding of the cognitive processes underlying attentional and interpretation biases in anxiety and depression. The third and final aim was to compare the test retest reliability of the parameters returned by fast-dm analysis against the attentional bias score to determine whether fast-dm has the capacity to be a more sensitive, and reliable, alternative means of analysis.

Assessing Model Fit

For each of the measures used in this thesis, the diffusion model appeared to fit the data well. However, it must be cautioned that as a non-mathematical psychologist, the fits were assessed based on the fit statistic returned by fast-dm 30, in combination with graphical inspection as recommended by Voss, Voss, & Lerche (2015). The models chosen for the dot probe and yes/no tasks were based on a theoretical understanding of the processes most likely to be responsible for attention and interpretation bias.

At the outset, the hypothesis was that attentional bias could be the result of three implicit decisional processes. First, individuals who show an attentional bias for disorder-specific negative emotion may have a lower threshold for responding to threats and dysphoric stimuli than those who do not show such a bias. Second, it may be that evidence for negative stimuli accumulates more quickly for individuals who show an attentional bias for negative emotion than those who do not show such a bias. Or, third, individuals who show an attentional bias for negative stimuli may have faster motor execution for processing negative items than individuals who do not show such a bias, which would result in shorter non-decision times. Therefore, we were looking for differences in starting point, drift rate, and non-decision time.

There were two ways we could map the data from the dot probe task to the diffusion model. In most circumstances, the model is mapped to the response keys. The response keys for the studies in this thesis were either left/right, or up/down, depending on the study. Therefore, the upper and

lower response thresholds represent each of the available response choices (i.e., left/right, up/down). Mapping the data in this manner means the parameters are based on evidence for the probe appearing on the left or right, or for the arrow pointing up and down. To examine the impact of affective stimuli, parameters would be compared based on trial type (congruent/incongruent) and emotion (happy/sad/angry). Differences based on trial type and emotion would then offer insight into how the emotive stimuli impacted the implicit decision-making parameters separated by the diffusion model.

An alternative way to map data from the dot probe task to the diffusion model is to map the response thresholds by congruence. For example, congruent responses are represented by the upper threshold, and incongruent responses are represented by the lower threshold. This is the way in which the parameters were mapped in the research in this thesis. Mapping the data this way enables direct feedback of the impact of trial type (congruent/incongruent), such that the parameters directly reflect bias toward or away from congruent stimulus/dot pairings. This is based on the premise that if attentional bias is consistent, when negative information is displayed, attention should be captured by the negative information, and thus, the equivalent processes in the opposite direction for the other stimulus should shift also (Rodebaugh et al., 2016).

Mapping the parameters to the task in this format means that boundary separation (*a*) represents how much evidence is required of the existence of the probe before responding, and start point (z_r) represents an expectation for the probe to appear more on congruent or incongruent trials. Drift rate (*v*) represents the rate at which evidence accumulates for congruent or incongruent responses. If evidence accumulates more quickly for congruent responses, this will result in a positive drift rate, whereas if evidence accumulates more quickly for incongruent responses, this results in a negative drift rate. As non-decisional times (*t0*) represent the speed of processes outside the decisional components of responding, they are theoretically not affected by how the parameters are mapped.

Choosing which parameters to allow to vary within the model is another theoretical consideration. The way the dot probe task is constructed in the current studies, different affective stimuli are presented intermixed in the same block, and there are equal numbers of congruent and incongruent trials, presented randomly within each block of the task. Therefore, there was no emotional priming prior to stimulus presentation. As such, there was no theoretical reason to expect that participants would adopt different boundary separation or starting point values for the existence of the probe from trial to trial, based on the valence of the emotion presented. However, there are individual differences in caution and decision bias. Some people require more evidence to execute a decision, whereas others need less evidence, and some individuals have a subconscious expectation for emotional (congruent trials) relative to neutral information (incongruent trials). For these reasons, boundary separation and starting point were permitted to vary by participant, but not emotion. Drift rate (*v*) and non-decision time (*t0*), however, are direct responses to the stimulus presented on each trial. As such, they may differ depending on the valence of the expression. One expression may elicit a stronger amount of evidence than others, which would present in drift rate. Alternatively, a specific expression may evoke a phasic response, which would present in non-decision time. As such, drift rate and non-decision time were free to vary between conditions and participants.

Turning now to the yes/no task, the data were mapped according to response: "Yes" was mapped to the upper boundary, "no" was mapped to the lower boundary. Therefore, boundary separation (*a*) represents how much evidence is required for the existence of the target emotion before responding. Start point (z_r) represents an expectation for the target emotion to appear more frequently than the distractor emotion. Drift rate (*v*) represents the rate at which evidence accumulates for the target or distractor emotions. A positive drift rate represents evidence accumulating more quickly for the target emotion; conversely, a negative drift rate results from evidence accumulating more quickly for distractor responses.

All four parameters were free to vary across participants, and trials. In the yes/no task, stimuli are presented in blocks, with a question to prime the expectation of the target valence for that block, e.g., Are these faces happy? This priming meant that both decision caution and decisional bias could vary depending on target expression in addition to drift rate and non-decision time. For example, an individual may require less evidence to interpret a stimulus as happy than as sad. As such, the boundary separation parameter would be smaller for the happy condition than the sad condition. In addition, an individual with a decisional bias for angry emotion would expect that angry expressions occur more frequently, and thus would have a starting point bias for the angry condition, but not the alternative conditions. Therefore, for the yes/no task, all parameters were free to vary by emotional condition, and between participants. Each condition was run through the model separately, as recommended by Voss, Voss, and Lerche (2015).

The final step was assessing model fit. This is one of the more challenging aspects of using evidence accumulation models for the non-specialist. It is laborious and time consuming. Using the cumulative density function that accompanies the fast-dm 30 package, predicted data is obtained from the parameters returned from the fast-dm 30 analysis. These return a set of predicted RTs that are plotted against the empirical data to inspect visually. If the predicted data align closely with the actual data, the fits are considered good. For the dot probe task, the fits appeared to be very good. For the yes/no task, there was some expectation by the predicted data that RTs at the lower end of the RT distribution would be faster than the empirical data demonstrated, however the response times at the upper end of the RT distribution appeared to fit well.

One of the biggest challenges to model fitting is finding the balance between theoretical premise and how well the model replicates the data. If a model replicates the data perfectly, but does not fit with any theoretical premise, then it is as limited as trying to use a model that does not replicate the data well. For this reason, when first fitting the model, basing it on theoretical premise, followed by graphical inspection of the resultant fit indices, is advised (Voss, Voss & Lerche, 2015). From there, alternative models can be examined, if necessary.

This thesis presents research that, to my knowledge, has for the first time fit the diffusion decision model to data from the dot probe task. As the graphical inspection of the fits look good, the first aim of this thesis was achieved, namely successful fitting of the dot probe task. Despite the challenges that can arise from the complexity of model fitting, the parameters of the diffusion model have provided more detail than RTs, which benefits our understanding of the cognitive processes that underlie implicit decision making, and are, arguably, worth the effort.

Cognitive Processes Underlying Attention and Interpretation in Anxiety and Depression

While this thesis was unable to answer the specific question of which processes are responsible for attentional bias and interpretation bias based on the study findings, it did identify some differences in the implicit decisional processes associated with anxiety and depression.

Attentional bias. Diffusion model analysis identified that both anxiety and depression are associated with a processing bias for emotional information. In Study 3, this was identified through faster accumulation of evidence for the probe when it is behind emotive information than when it is behind neutral information for individuals with a history of anxiety and depression. This pattern was not found in individuals with no such history. Additionally, individuals with a history of depression demonstrated a phasic response to emotional stimuli, which was evident from faster non-decision times compared to individuals with a non-clinical history. This was not extant for individuals with anxiety.

A processing bias for emotional information was also identified in Study 1. This was initially attributed to a non-clinical population, as there was no disorder-congruent preference for one emotion over another, which we had hypothesised. However, considering the findings from Study 3, further inspection of the results from Study 1 identified that the mean scores on the DASS and DERS were higher for the non-clinical sample recruited in Study 1 than they were for the participants with a history of anxiety or depression in Study 3. Study 1 also found that drift rate was related to increased anxiety and depression, while slower non-decisional processes and more cautious responding were a protective factor to emotion regulation difficulties. Therefore, taking the

between-groups differences in processing from Study 3, which identified a processing bias for emotion only in individuals with a clinical history, with the correlational analyses from Study 1, the findings suggest that implicit engagement with emotion may be a key underlying factor in anxiety and depression.

Theoretically, these findings are in line with Beck's (1967) hypothesis that attentional capture by emotional information is linked with affective disorders. However, in contrast to Beck's theory, our findings are not content-specific. Beck's content specificity hypothesis postulates that individuals with anxiety and depression attend to disorder-congruent stimuli in the environment. Specifically, those with anxiety attend to threat, and those with depression notice dysphoric information more saliently than individuals without these disorders. In the studies presented in this thesis, there were no disorder-congruent attentional biases identified. However, as outlined above, there does seem to be an attentional engagement with emotive stimuli in general.

Attentional bias to disorder-congruent stimuli was initially theorised to be a stable trait in individuals with affective disorders (see Mogg & Bradley, 2016, for a review of the theoretical accounts of attentional bias). However, more and more research is examining the possibility of the dynamic nature of attentional biases. The fact that we did not find disorder-congruent attentional biases in Study 3, even in those who have high levels of anxiety and depression symptoms, provides support for the proposition that attentional bias may be dynamic in nature (e.g. (Zvielli et al., 2015).

Interpretation bias. Diffusion model analysis identified that interpreting more ambiguous information impacts rate of evidence accumulation (drift rate) and decision bias (starting point). Further, individuals with a history of anxiety have a propensity to accumulate evidence more quickly across all emotions when interpreting the category of an emotion than individuals without a history of anxiety. This may be a function of hypervigilance in that they are quicker to analyse and interpret the stimuli in their environment due to a heightened fear of threat. In contrast, individuals with a history of depression demonstrated a more conservative decision criterion in their responding

than individuals with no history of depression. Thus, individuals with depression required more evidence to classify emotional information.

The broader implications of these findings are that there do appear to be different, and specific, processes responsible for implicit decision making in anxiety and depression. Therefore, the primary benefit of using the diffusion decision model in this area of research is that it can identify these specific decisional processes, and broaden our understanding of how decisional processes differ between individuals with affective disorders and those without. This broader understanding enables the development of interventions targeted at the specific underlying processes, with the diffusion model providing the capacity to assess the efficacy of such interventions. Further, by using the diffusion model to analyse existing data that *have* been successful in capturing attentional bias through RTs, the specific processes of change can be identified. This will help identify individuals whose biases are a result of those same processes that are targeted by ABM, and thus for whom ABM will be most effective.

Test-Retest Reliability of the Diffusion Model Parameters

The final aim of this thesis was to compare the test-retest reliability of the parameters returned by fast-dm analysis against the attentional bias score. The goal was to assess whether fast-dm has the capacity to be a more sensitive, reliable alternative. While the results throughout this thesis demonstrate the increased sensitivity of diffusion model analysis relative to RT analysis, the test-retest reliability of the parameters derived from fast-dm-30 was mixed. There was no correlation between the traditional attentional bias score at pre-training and post training, for any valence. For the fast-dm parameters, the pre- and post-training values for starting point and drift rate did not correlate. However, non-decision time values from pre- and post-training were highly correlated across time for all emotions, and boundary separation demonstrated moderate test-retest reliability. While these results do not offer unequivocal support for the test-retest reliability of fast-dm parameters, there are some methodological factors to consider before concluding that these reliability coefficients may mean the model is not suitable.

In particular, lack of variability in data points can reduce the likelihood of a correlation (Goodwin & Leech, 2006). The drift rate values throughout the studies in this thesis had minimal variability. This may be a function of the lack of perceptual bias identified generally, or it may be a function of how the dot probe task was mapped to the diffusion model. I chose to map the model so that the drift rate represents a difference in evidence accumulation for congruent and incongruent trials. It may be that mapping the model to the responses (as outlined above) would offer more variability, and thus better test-retest stability. Alternatively, it may be that attentional bias is not a function of perceptual bias, and thus is not captured by drift rate, or that the parameters of the diffusion model, specifically the drift rate, are not a more reliable way to assess attentional bias.

One way to assess these possibilities is to apply fast-dm 30 to data derived from an existing dataset that has already found an attentional bias with RT at two time-points. If there is no capture of the bias by drift rate, we can hypothesise that attentional bias is not a function of perceptual bias. If attentional bias is captured by drift rate, then correlating the values of drift rate at two time-points will identify its test-retest reliability. If so, then the cognitive processes that are modified by ABM can be explored.

Summary

This thesis has presented the successful application of diffusion decision model analysis to data from the dot probe task for the first time. In doing so, the research presented has demonstrated the capacity for the diffusion model to identify implicit decisional processes that differ between anxiety and depression that are not captured by RTs alone. While the test-retest reliability of the fast-dm 30 parameters was mixed, guidance has been provided for future research to gain clarity on the reliability of the diffusion model parameters, and their suitability as an alternative measure of attentional bias. Finally, the research in this thesis has demonstrated the value of adopting a mathematical psychology analytical approach to the field of applied clinical psychology. Specifically, the ability to identify differences in decisional processing between anxiety and depression provides a platform for research to more effectively identify and target the processes that

underlie clinical disorders. By doing so, the benefits to clinical psychology are increased awareness of causal factors, and enhanced treatment options, and hopefully, outcomes.

References

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders-DSM-5. *Washington, DC: Author*.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. International Universities Press.

Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016). Scaling-up treatment of depression and anxiety: a global return on investment analysis. *The Lancet. Psychiatry*, *3*(5), 415–424. https://doi.org/10.1016/S2215-0366(16)30024-4

Goodwin, L. D., & Leech, N. L. (2006). Understanding Correlation: Factors That Affect the Size of r. *The Journal of Experimental Education*, *74*(3), 249–266. https://doi.org/10.3200/JEXE.74.3.249-266

Heathcote, A., Suraev, A., Curley, S., Gong, Q., Love, J., & Michie, P. T. (2015). *Decision processes and the slowing of simple choices in schizophrenia.* American Psychological Association. Retrieved from http://psycnet.apa.org/journals/abn/124/4/961/

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, 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. https://doi.org/10.1037//0021-843x.111.1.107

Mogg, K., & Bradley, B. P. (2016). Anxiety and attention to threat: Cognitive mechanisms and treatment with attention bias modification. *Behaviour Research and Therapy*, *87*(Supplement C), 76–108. https://doi.org/10.1016/j.brat.2016.08.001

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59.

Renaud, J., Russell, J. J., & Myhr, G. (2014). Predicting Who Benefits Most From Cognitive-Behavioral Therapy for Anxiety and Depression. *Journal of Clinical Psychology*, *70*(10), 924–932. https://doi.org/10.1002/jclp.22099

Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods*, *39*(4), 767–775. https://doi.org/10.3758/BF03192967

Voss, A., Voss, J., & Lerche, V. (2015). Assessing cognitive processes with diffusion model analyses: a tutorial based on fast-dm-30. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00336

Vandekerckhove, J., & Tuerlinckx, F. (2008). Diffusion model analysis with MATLAB: A DMAT primer. *Behavior Research Methods*, *40*(1), 61–72. https://doi.org/10.3758/BRM.40.1.61

Wagenmakers, E.-J., van der Maas, H. L. J., & Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*(1), 3–22.

White, C. N., Ratcliff, R., Vasey, M. W., & McKoon, G. (2010). Using diffusion models to understand clinical disorders. *J Math Psychol*, *54*, 39–52. https://doi.org/10.1016/j.jmp.2010.01.004

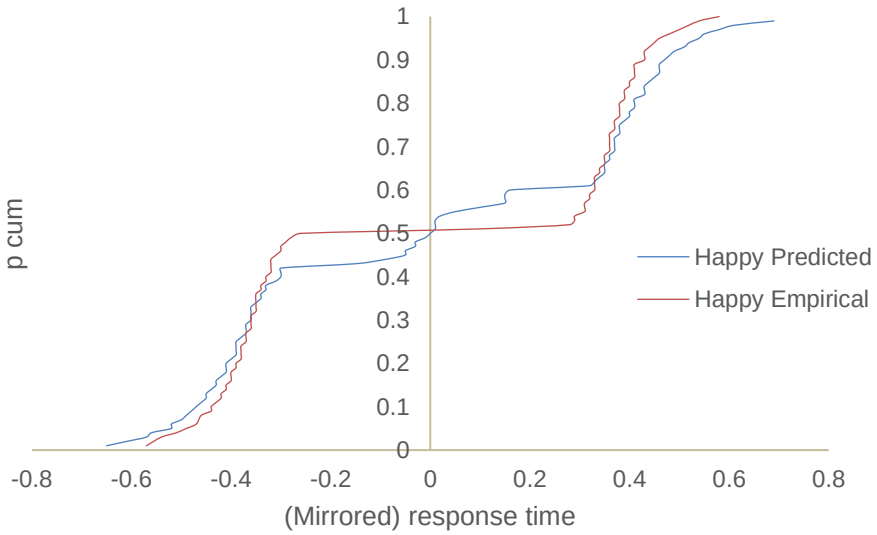
White, Corey N., Skokin, K., Carlos, B., & Weaver, A. (2016). Using decision models to decompose anxiety-related bias in threat classification. *Emotion*, *16*(2), 196.

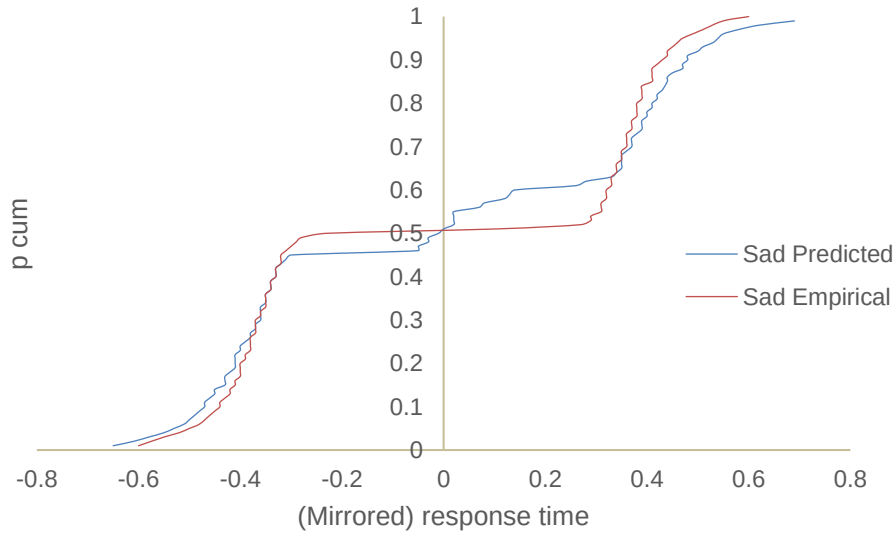
Zvielli, A., Bernstein, A., & Koster, E. H. (2015). Temporal dynamics of attentional bias. *Clinical Psychological Science*, *3*(5), 772–788.

Appendix 1

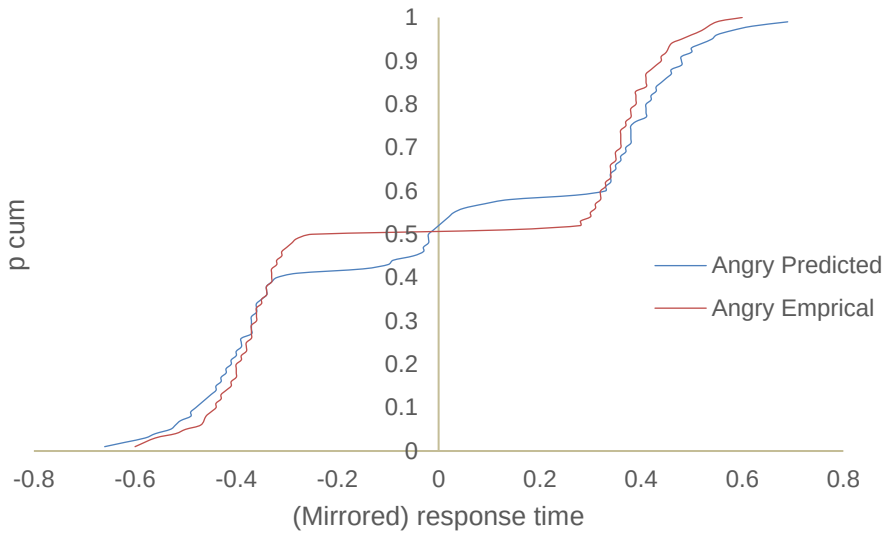
Study 1 Supplementary material

Comparison of empirical and predicted response time distributions for each expression (happy, sad, angry) are shown below. Times for congruent trials are represented in the right side of the graph, incongruent trials have been mirrored on the zero point of the time axis, represented in the left part of the graph. The darker line is the accumulated probability function computed according to the diffusion model. The lighter line shows the cumulative probability of empirical response times. Note that both cumulative functions must converge to 1. The flat portion of the line at 0.5 *P* cum is an absence of response times, as response times were generally slower than 200ms.

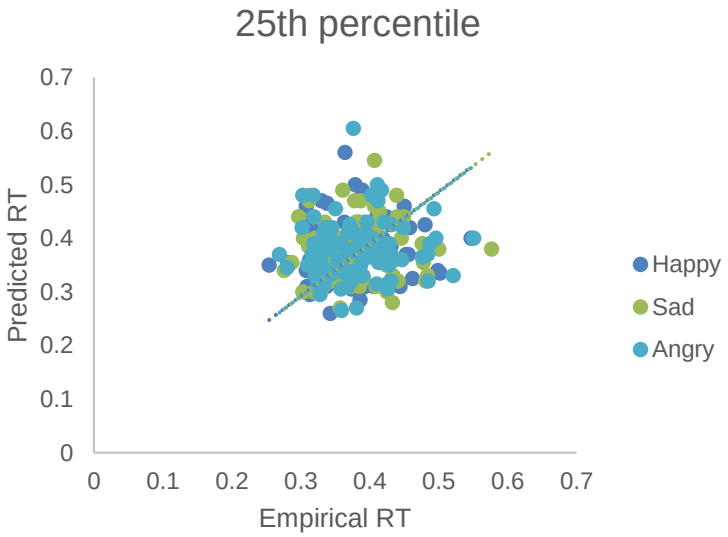


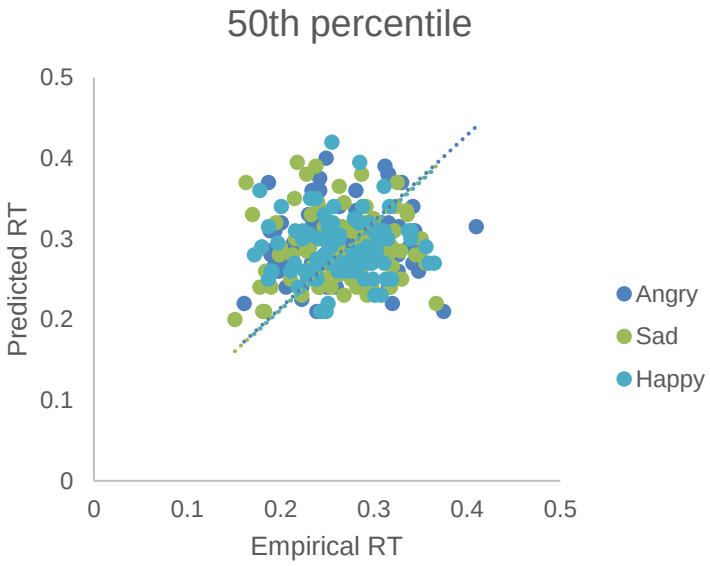


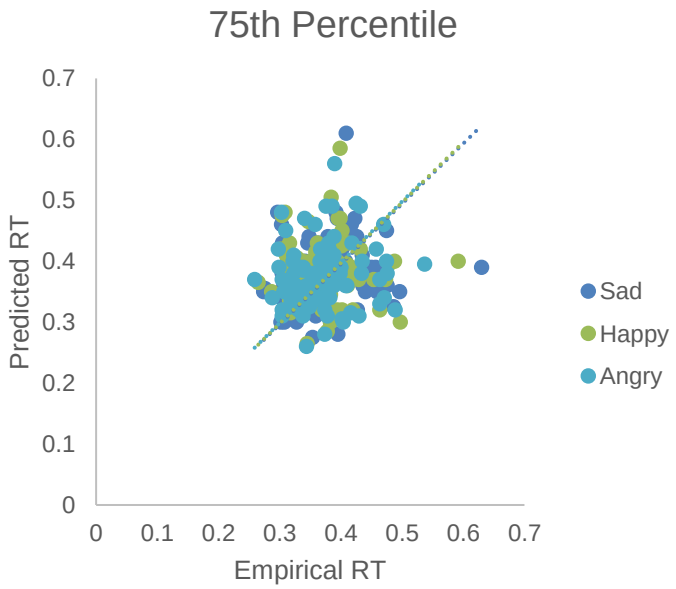
a



Quantile probability plots comparing empirical and predicted response times for the 25th, 50th, and 75th percentiles of each expression (happy, sad, angry) are presented below.



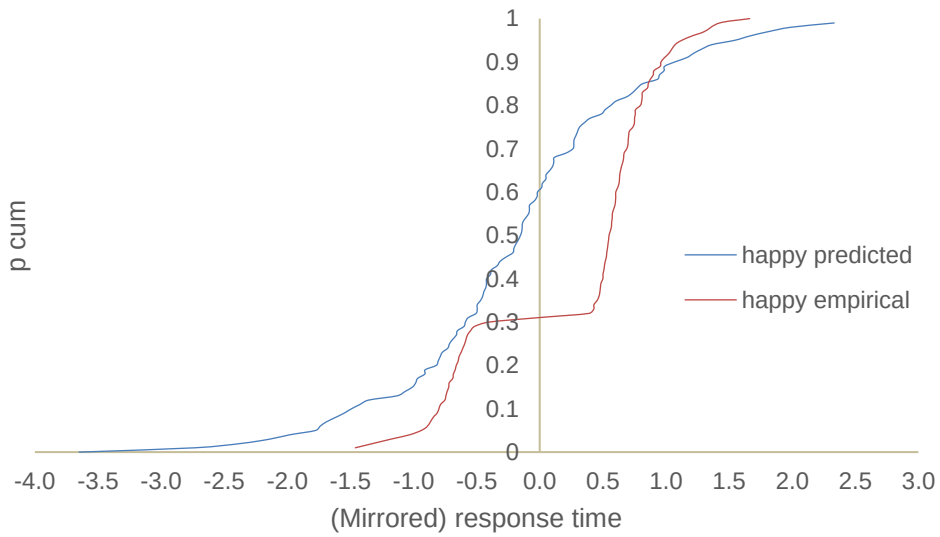

50th percentile
Predicted RT
Empirical RT
0
0.1
0.2
0.3
0.4
0.5
Angry
Sad
Happy

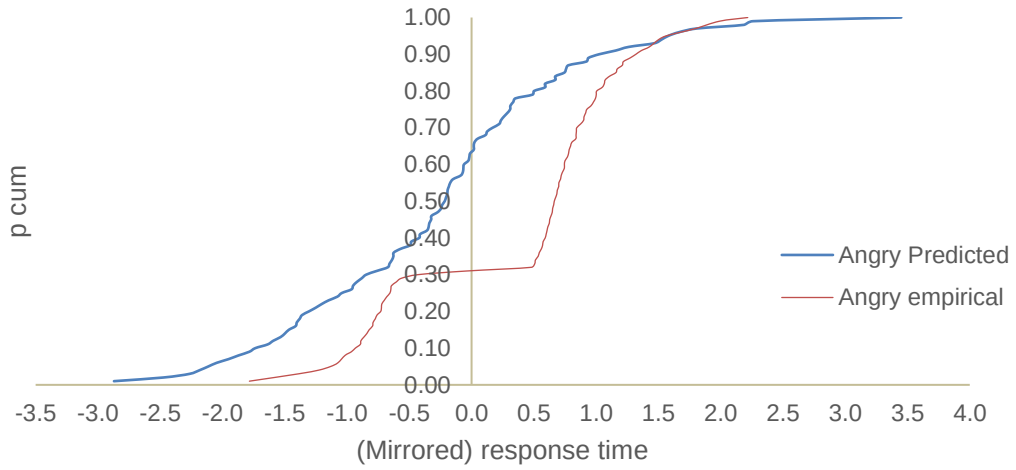

75th Percentile
Predicted RT
Empirical RT
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Sad
Happy
Angry

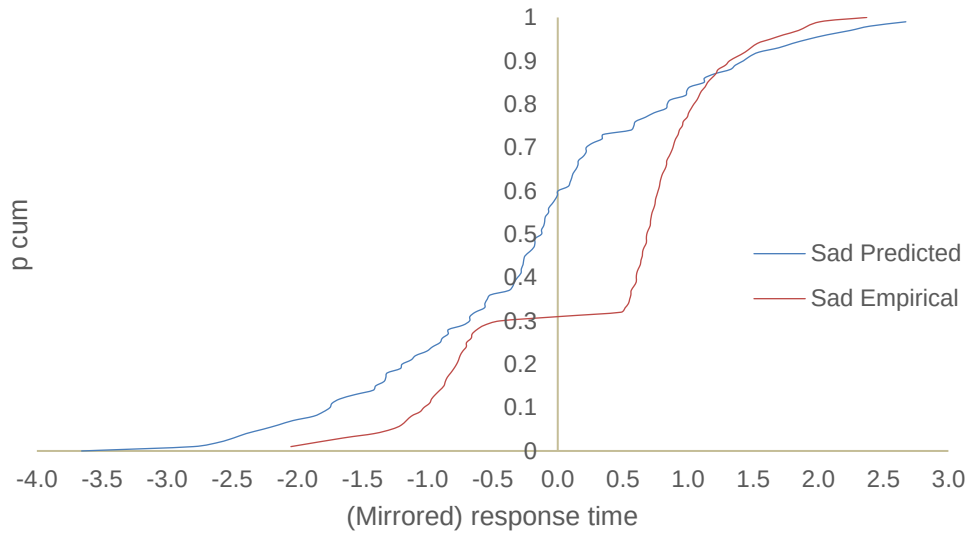
Appendix 2

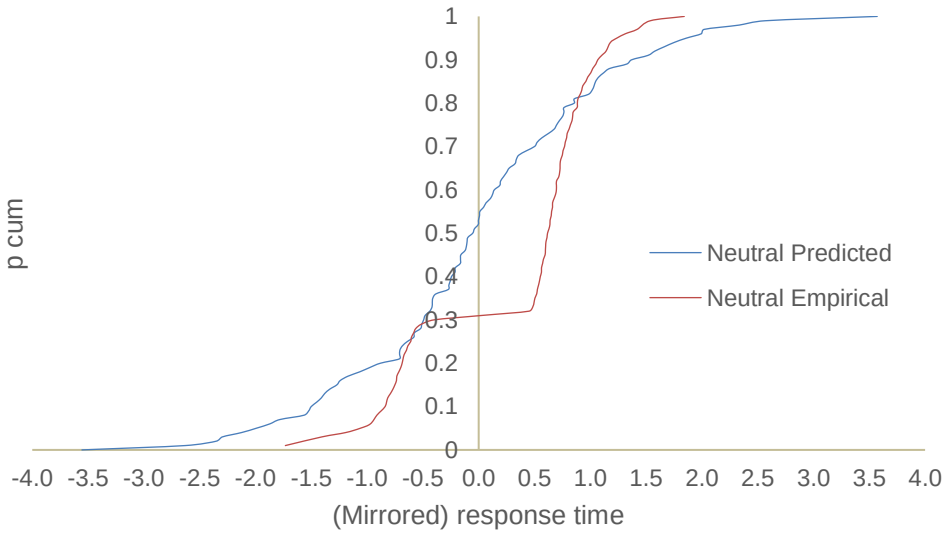
Study 2 Supplementary material

Comparison of empirical and predicted response time distributions for each expression (happy, angry, sad, neutral) are shown below. Times for "Yes" responses are represented in the right side of the graph, "No" responses have been mirrored on the zero point of the time axis, represented in the left part of the graph. The darker line is the accumulated probability function computed according to the diffusion model. The lighter line shows the cumulative probability of empirical response times. Note that both cumulative functions must converge to 1. The flat portion of the line at 0.3 *P* cum is an absence of response times, as response times were generally slower than 500ms.



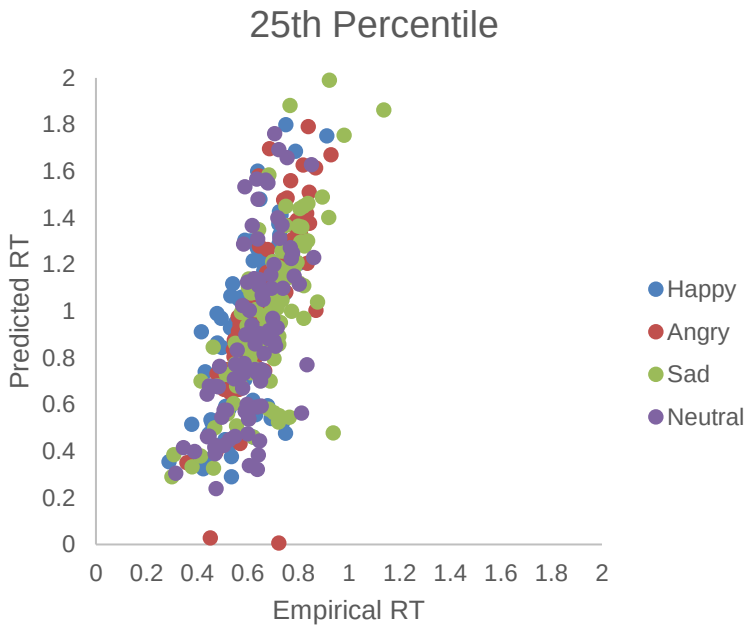

1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
p cum
-3.5 -3.0 -2.5 -2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0
(Mirrored) response time
Angry Predicted
Angry empirical

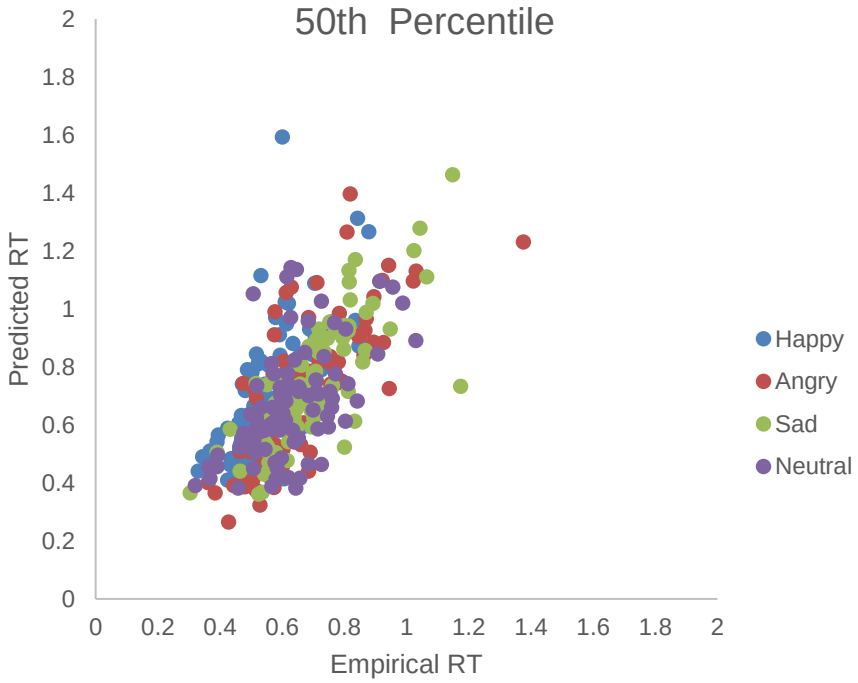

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
p cum
-4.0 -3.5 -3.0 -2.5 -2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0 2.5 3.0
(Mirrored) response time
Sad Predicted
Sad Empirical

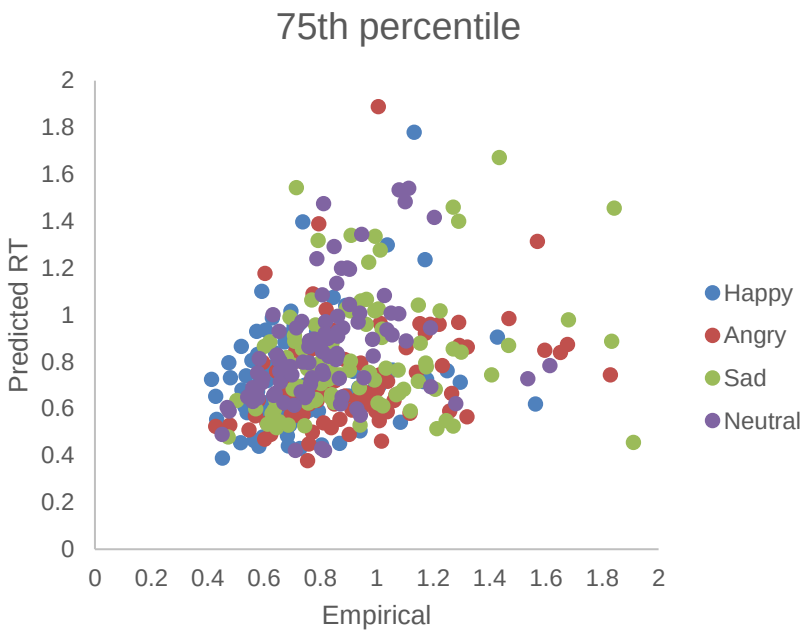

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
p cum
-4.0 -3.5 -3.0 -2.5 -2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0
(Mirrored) response time
Neutral Predicted
Neutral Empirical

e

Quantile probability plots comparing empirical and predicted response times for the 25th, 50th, and 75th percentiles of each expression (happy, sad, angry) are presented below.






75th percentile
Predicted RT
Empirical
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
Happy
Angry
Sad
Neutral

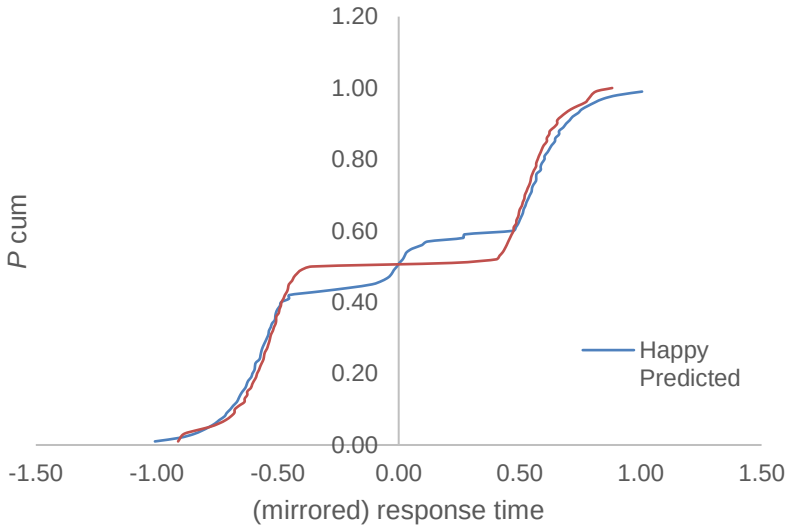
g

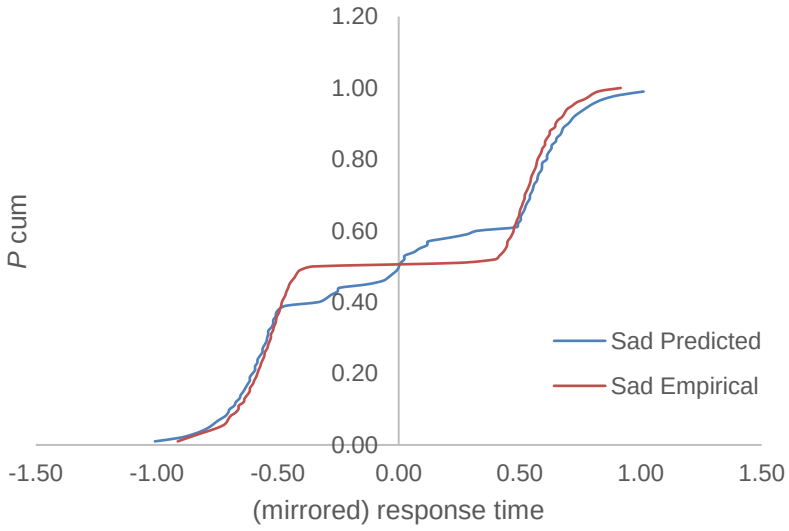
Appendix 3

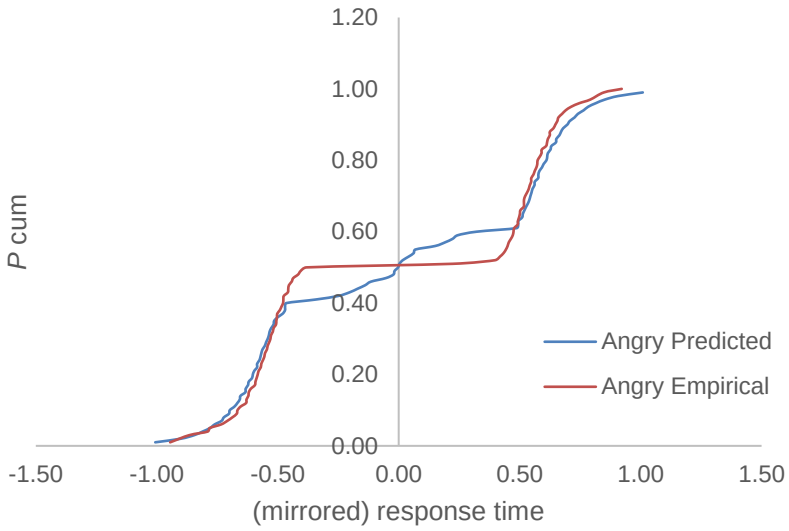
Study 3 Supplementary material

Dot Probe Pre-Training

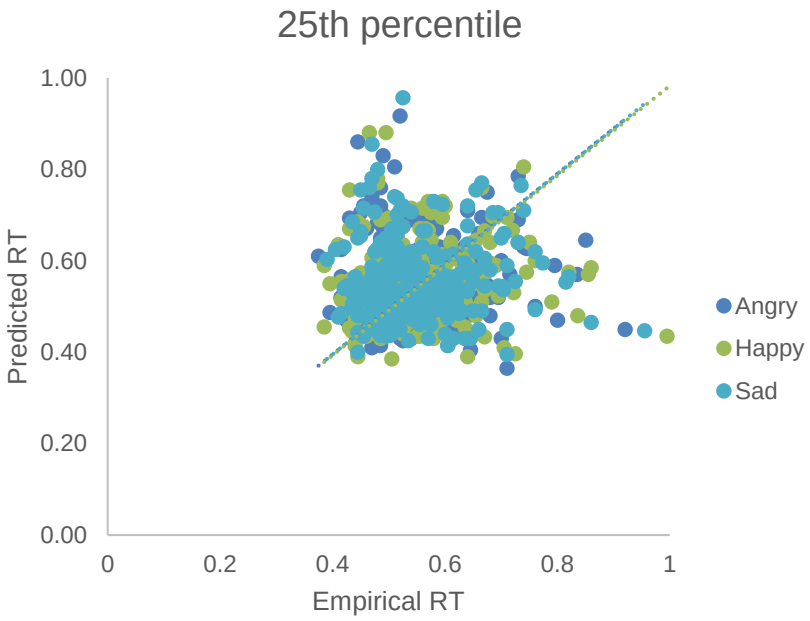
Comparison of empirical and predicted response time distributions for each expression (happy, sad, angry) are shown below. Times for congruent trials are represented in the right side of the graph, incongruent trials have been mirrored on the zero point of the time axis, represented in the left part of the graph. The darker line is the accumulated probability function computed according to the diffusion model. The lighter line shows the cumulative probability of empirical response times. Note that both cumulative functions must converge to 1. The flat portion of the line at 0.5 *P* cum is an absence of response times, as response times were generally slower than 200ms.

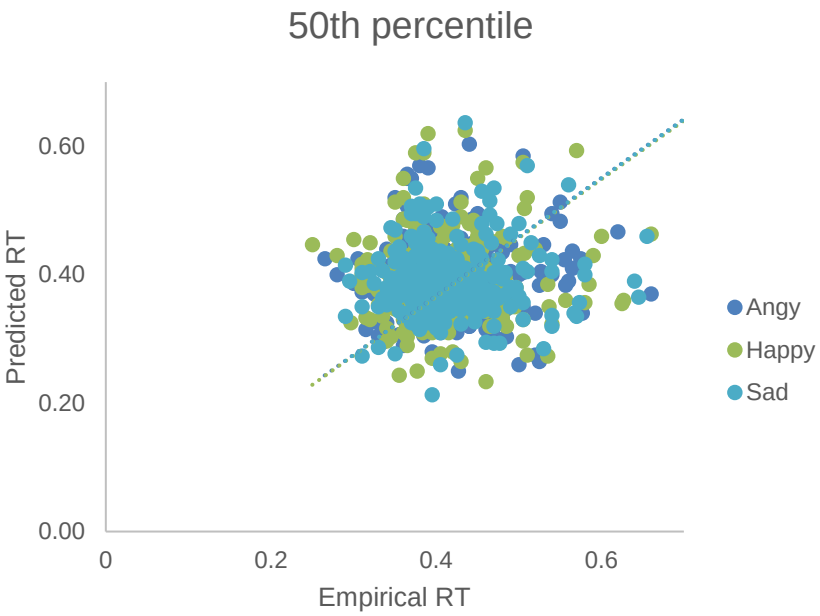


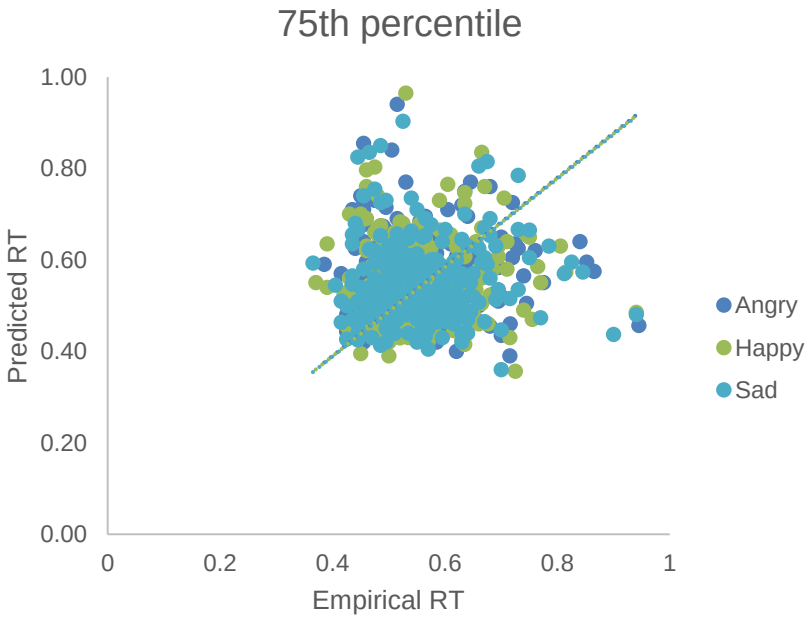




Quantile probability plots comparing empirical and predicted response times for the 25th, 50th, and 75th percentiles of each expression (happy, sad, angry) are presented below.

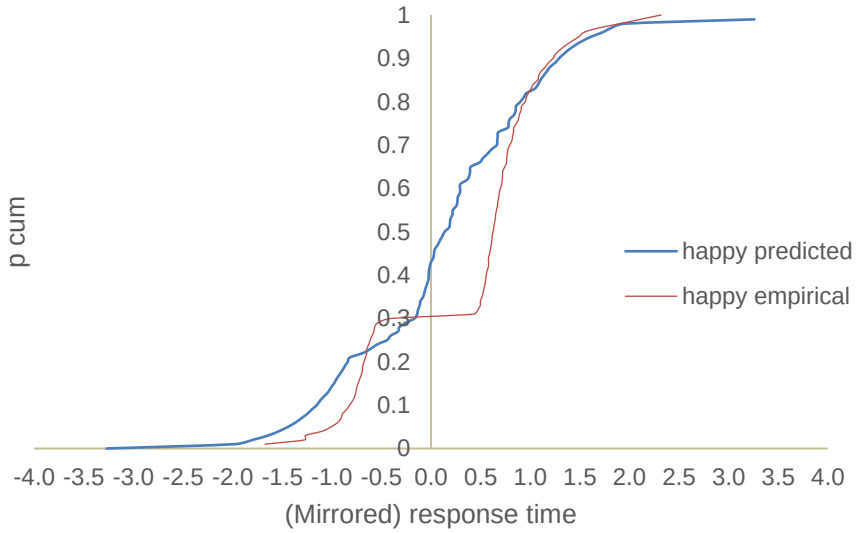


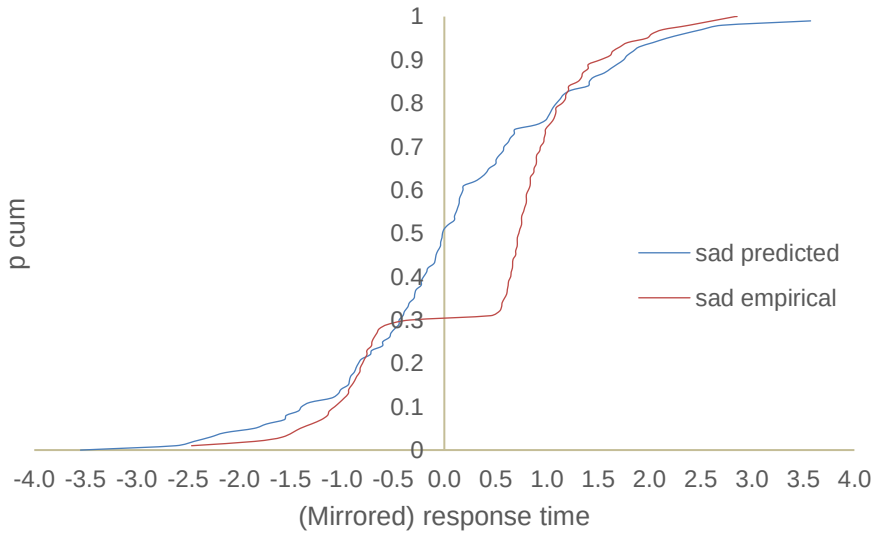

50th percentile
Predicted RT
0.60
0.40
0.20
0.00
0
0.2
0.4
0.6
Empirical RT
Angy
Happy
Sad

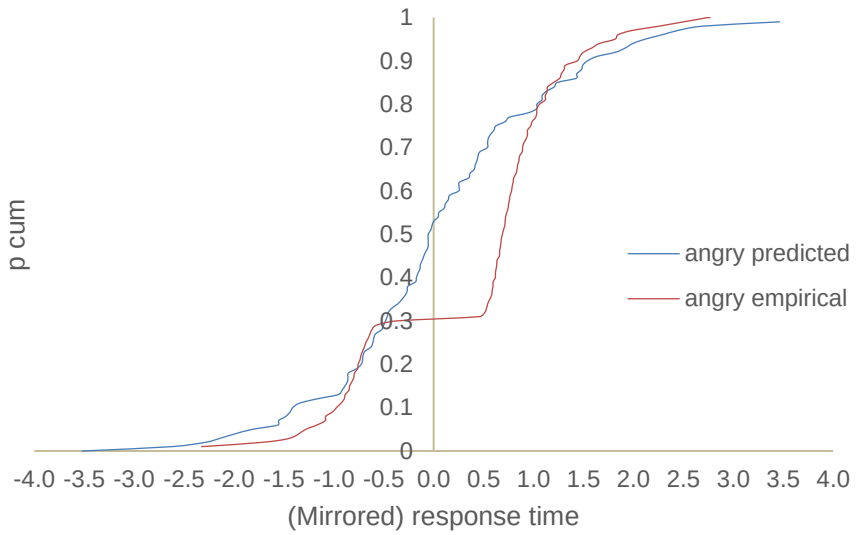

75th percentile
Predicted RT
1.00
0.80
0.60
0.40
0.20
0.00
0
0.2
0.4
0.6
0.8
1
Empirical RT
Angry
Happy
Sad

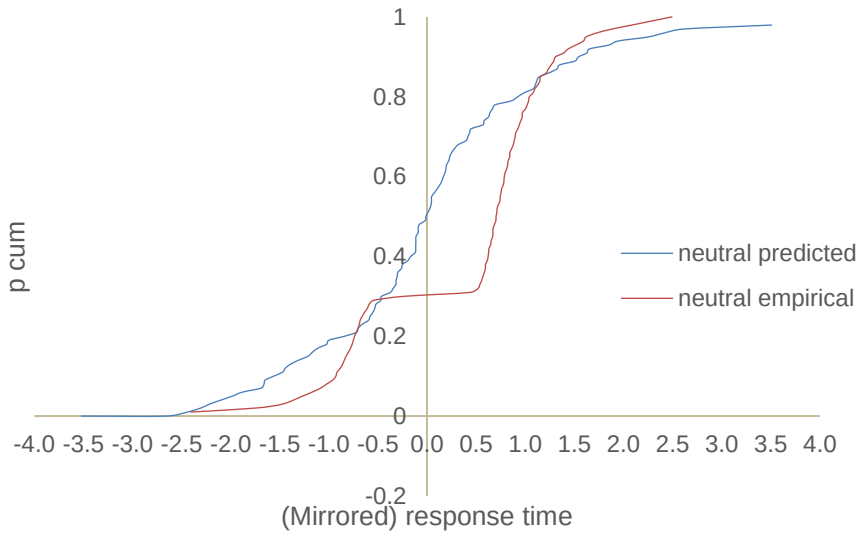
Yes/No Pre-training

Comparison of empirical and predicted response time distributions for each expression (happy, angry, sad, neutral) are shown below. Times for "Yes" responses are represented in the right side of the graph, "No" responses have been mirrored on the zero point of the time axis, represented in the left part of the graph. The darker line is the accumulated probability function computed according to the diffusion model. The lighter line shows the cumulative probability of empirical response times. Note that both cumulative functions must converge to 1. The flat portion of the line at 0.3 *P* cum is an absence of response times, as response times were generally slower than 500ms.

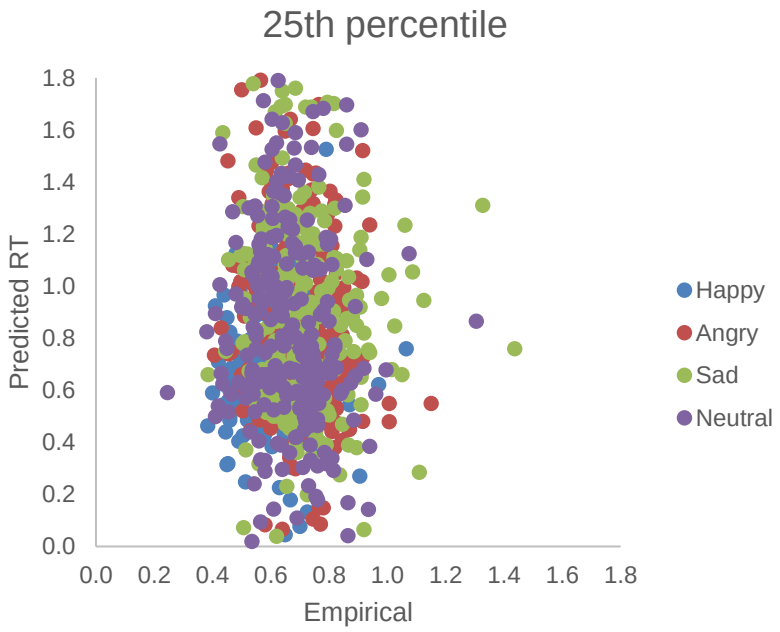


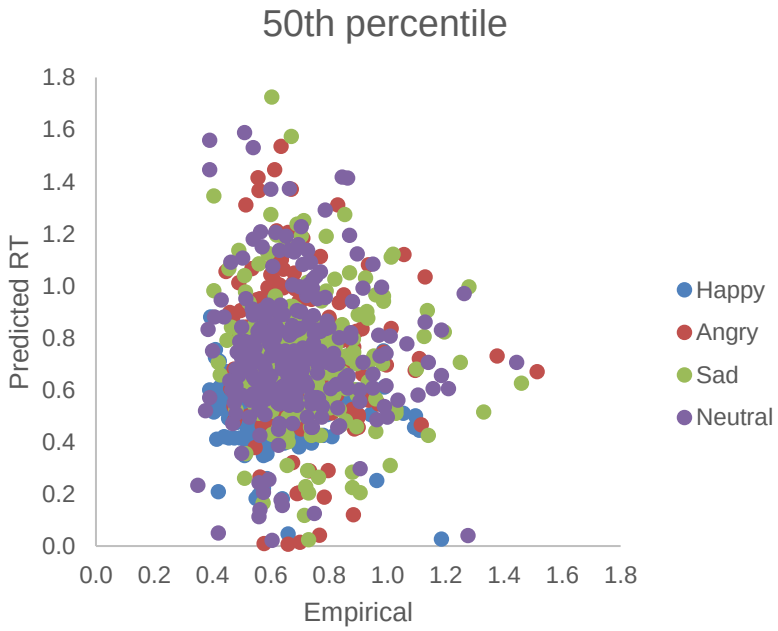



1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
p cum
angry predicted
angry empirical
-4.0 -3.5 -3.0 -2.5 -2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0
(Mirrored) response time

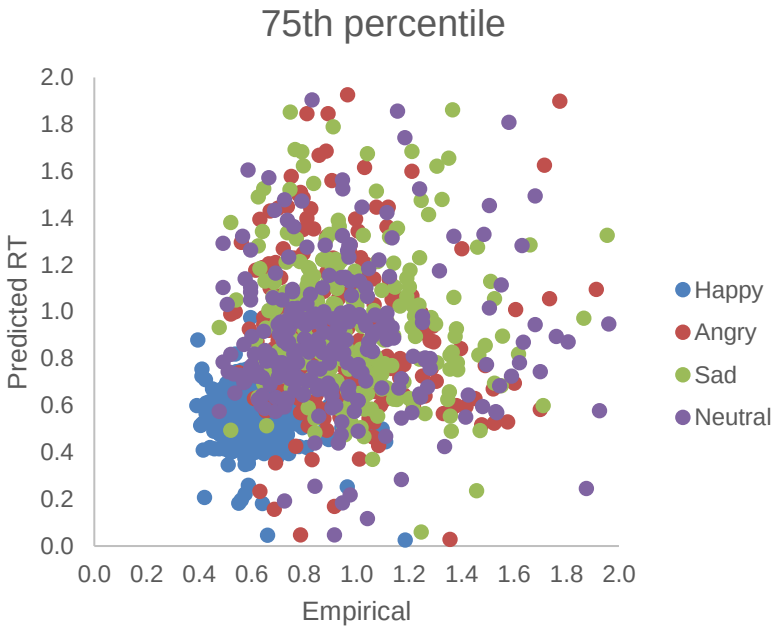

1
0.8
0.6
0.4
0.2
0
-0.2
p cum
neutral predicted
neutral empirical
-4.0 -3.5 -3.0 -2.5 -2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0
(Mirrored) response time

Quantile probability plots comparing empirical and predicted response times for the 25th, 50th, and 75th percentiles of each expression (happy, sad, angry) are presented below.



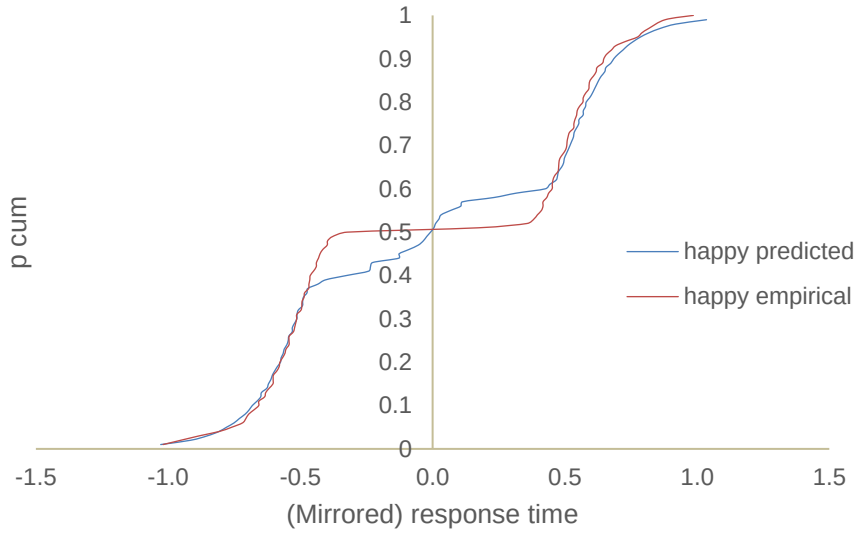


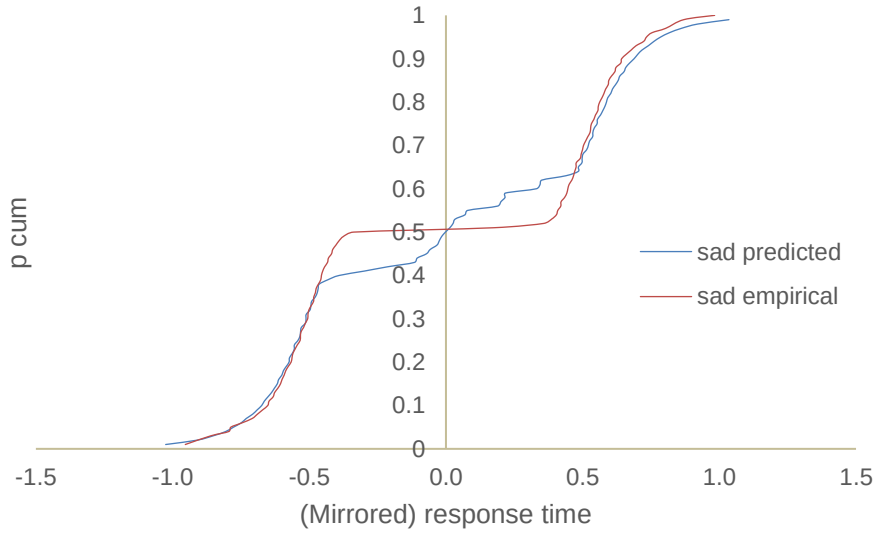
m

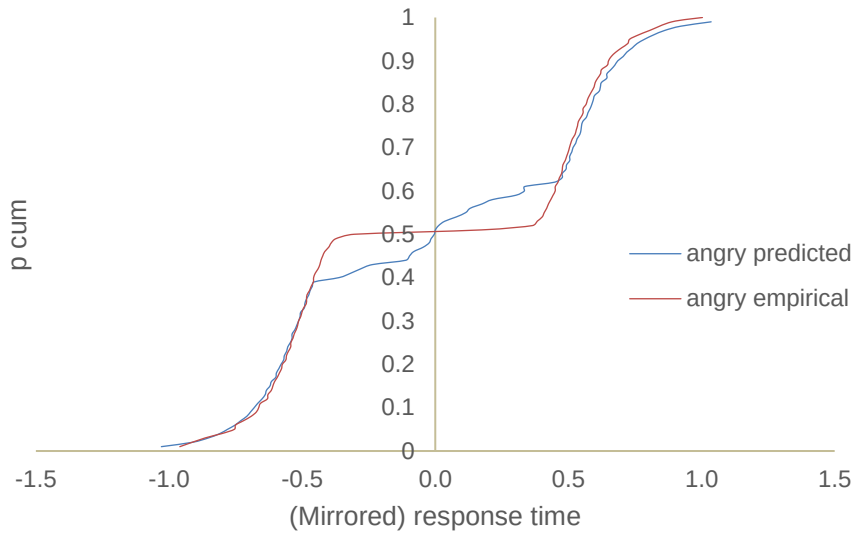

75th percentile
Predicted RT
Empirical
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
Happy
Angry
Sad
Neutral

Dot Probe Post-Training

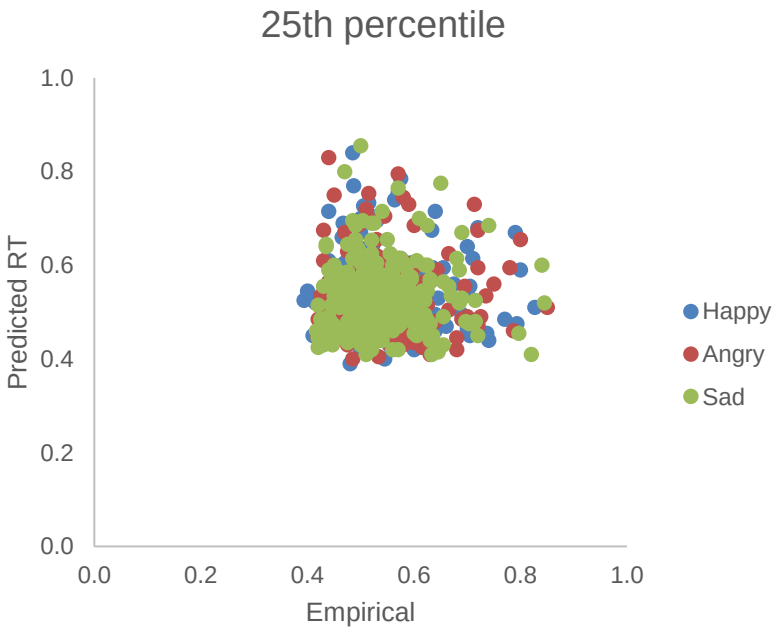
Comparison of empirical and predicted response time distributions for each expression (happy, sad, angry) are shown below. Times for congruent trials are represented in the right side of the graph, incongruent trials have been mirrored on the zero point of the time axis, represented in the left part of the graph. The darker line is the accumulated probability function computed according to the diffusion model. The lighter line shows the cumulative probability of empirical response times. Note that both cumulative functions must converge to 1. The flat portion of the line at 0.5 *P* cum is an absence of response times, as response times were generally slower than 200ms.

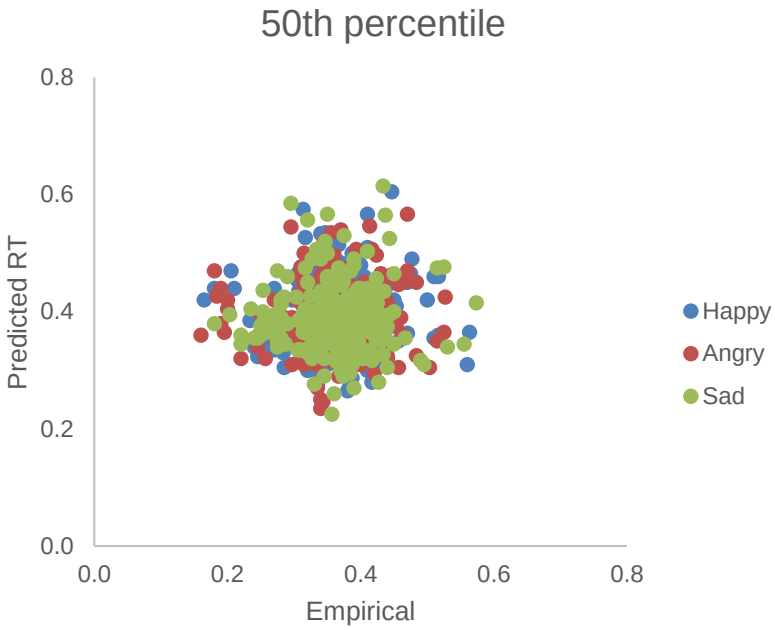


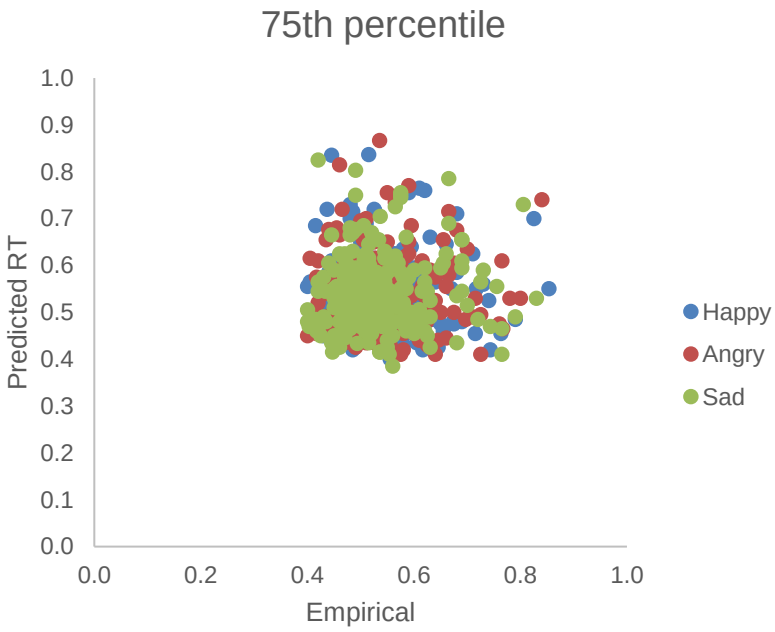




Quantile probability plots comparing empirical and predicted response times for the 25th, 50th, and 75th percentiles of each expression (happy, sad, angry) are presented below.

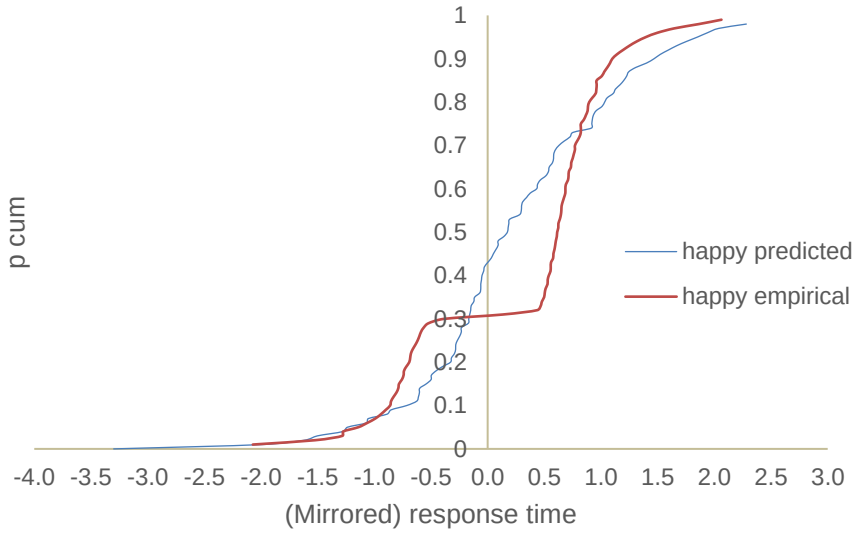


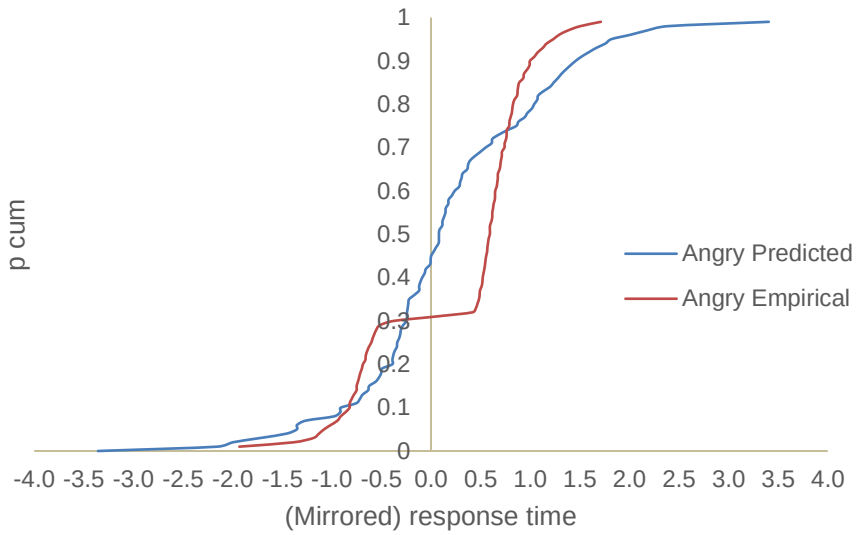

50th percentile
Predicted RT
0.0
0.2
0.4
0.6
0.8
Empirical
Happy
Angry
Sad

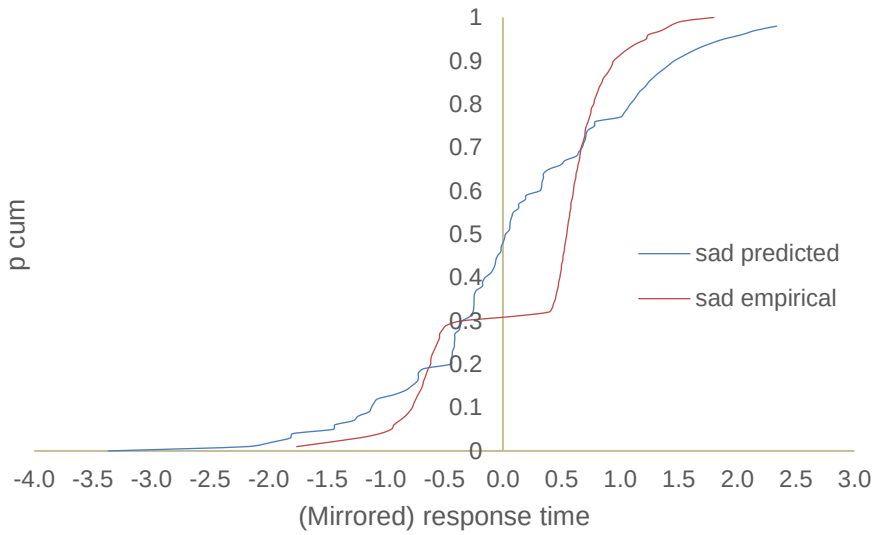

75th percentile
Predicted RT
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Empirical
Happy
Angry
Sad

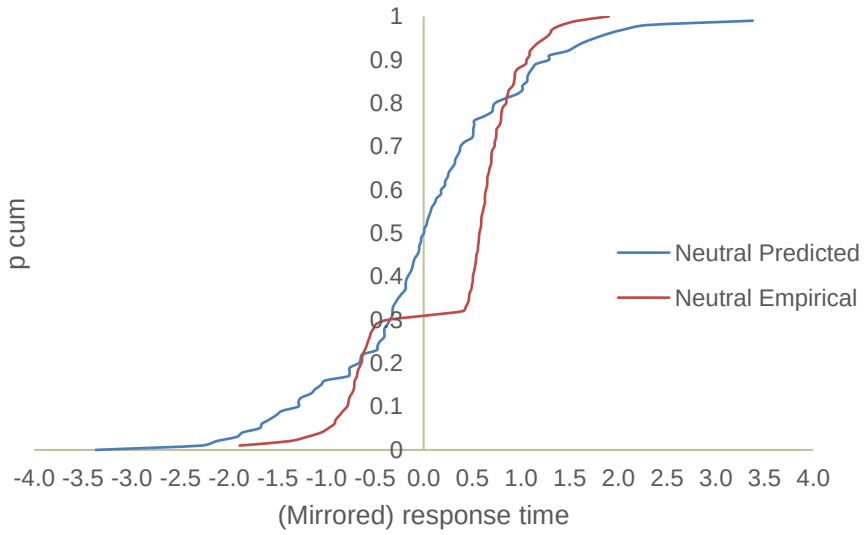
Yes/No Post Training

Comparison of empirical and predicted response time distributions for each expression (happy, angry, sad, neutral) are shown below. Times for "Yes" responses are represented in the right side of the graph, "No" responses have been mirrored on the zero point of the time axis, represented in the left part of the graph. The darker line is the accumulated probability function computed according to the diffusion model. The lighter line shows the cumulative probability of empirical response times. Note that both cumulative functions must converge to 1. The flat portion of the line at 0.3 *P* cum is an absence of response times, as response times were generally slower than 500ms.

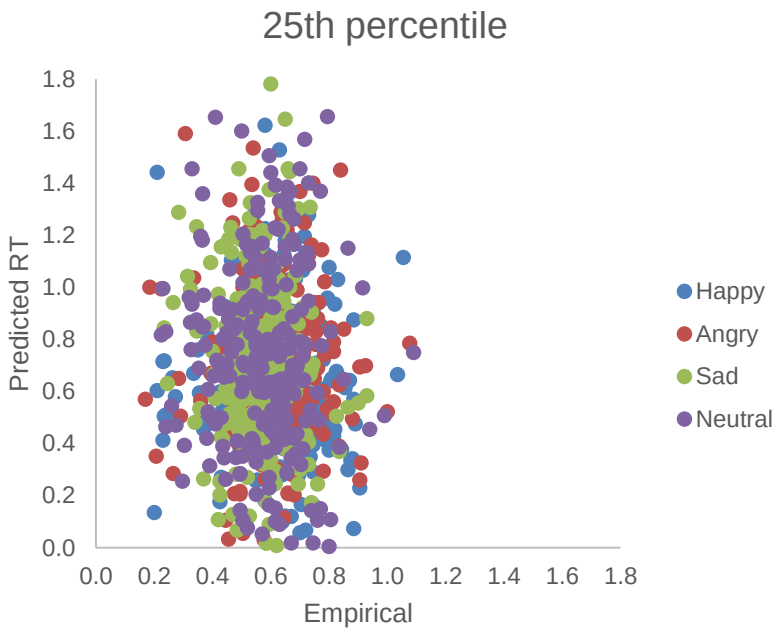


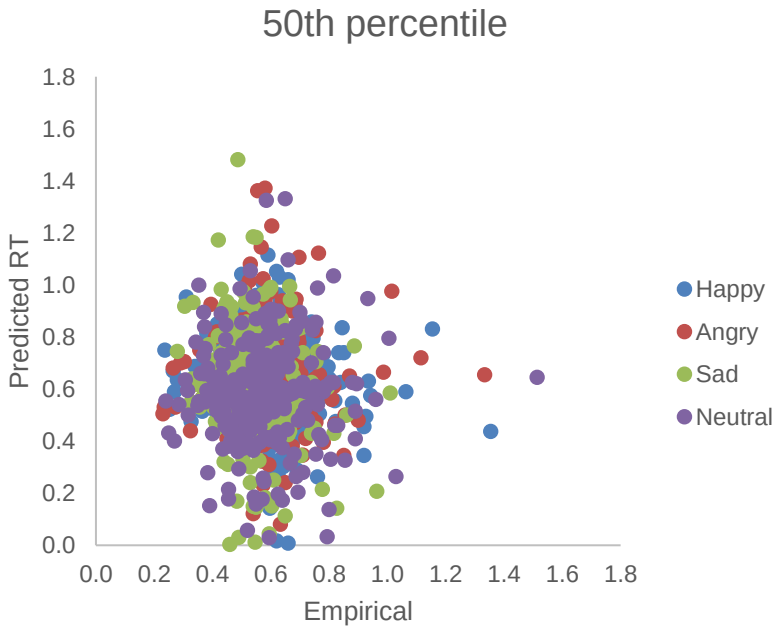

p cum
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-4.0 -3.5 -3.0 -2.5 -2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0
(Mirrored) response time
Angry Predicted
Angry Empirical


p cum
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-4.0 -3.5 -3.0 -2.5 -2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0 2.5 3.0
(Mirrored) response time
sad predicted
sad empirical

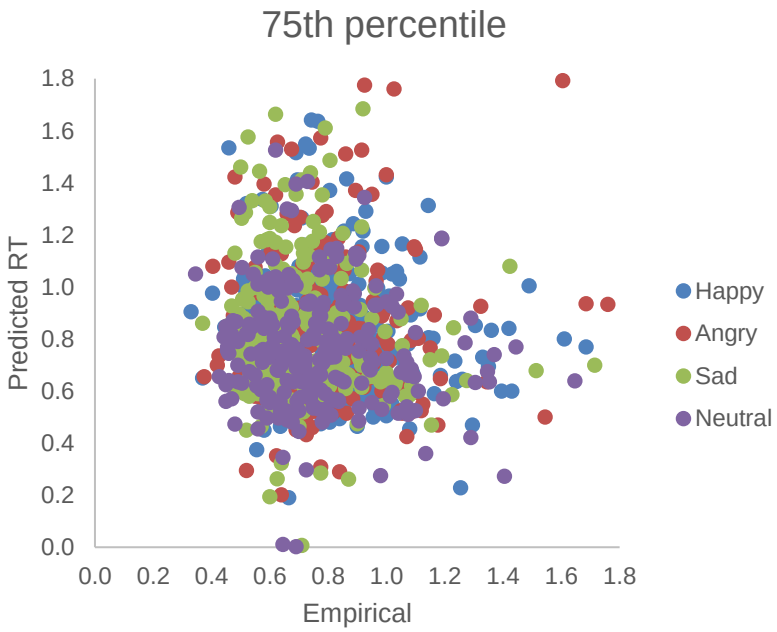

p cum
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-4.0 -3.5 -3.0 -2.5 -2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0
(Mirrored) response time
Neutral Predicted
Neutral Empirical

Quantile probability plots comparing empirical and predicted response times for the 25th, 50th, and 75th percentiles of each expression (happy, sad, angry) are presented below.





t


75th percentile
Predicted RT
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6 1.8
Empirical
Happy
Angry
Sad
Neutral