# The Mechanisms and Effects of Modifying Attentional Biases to Threatening Information

Michael Browning

St John's College

Thesis submitted for the degree of Doctor of Philosophy

Department of Psychiatry

The University of Oxford

Trinity Term, 2010

The Mechanisms and Effects of Modifying Attentional Biases to Threatening Information

Michael Browning, St John's College

Trinity Term 2010

# Abstract

Patients with both depression and anxiety show an increased tendency to deploy attention towards negative information. Cognitive models of the illnesses predict that these negative attentional biases are causally related to the symptoms of the disorders. Consistent with this, modifying attentional bias using either antidepressant medication or simple, computer based training tasks has previously been associated with altered symptomatology in both non-clinical and clinical populations. The current thesis aimed to investigate the mechanisms by which attentional bias training tasks alter attention. The investigations were conducted within an experimental neuroscience framework which has previously been successfully deployed in studies of antidepressant medication. The thesis then sought to use these initial results to improve the basic understanding of attentional control processes and, ultimately, guide the development of novel treatment strategies.

The initial studies of the thesis characterised the behavioural and neural effects of attentional bias training. Behaviourally, a high degree of generalisation of the training effect was found across a range of emotional stimuli. Neurally, training was found to alter activity in a network of prefrontal regions known to be involved in the control of attention. Further analysis, utilising a computational learning model, suggested that the attentional control systems identified in this study could be understood in terms of expectation based processes. These studies therefore indicated that, in contrast to the predominately limbic effects of antidepressant medication, training initially altered the response of frontal control circuitry.

The later studies of the thesis investigated possible strategies for extending the use of attentional bias training. Firstly, combining training with antidepressant medication was found to produce an interference effect on emotional memory suggesting that administering both interventions concurrently is likely to erode their cognitive impact. Lastly, attentional bias training was found not to alter attention in patients with bipolar disorder, with the results of the study indicating that standard assessments of attentional bias in this clinical population are likely to be unreliable. Overall, these studies indicate that attentional bias training may be used to alter the top-down control of attention to emotional information and suggest that such effects may interfere with the bottom-up effects of antidepressant drugs. More generally the work demonstrates the utility of using a cognitive-neuroscientific framework to explore the mechanisms and impact of novel therapeutic strategies.

# Contents

**Abstract** ii

**Acknowledgements** x

**Declaration** xi

**Abbreviations and Glossary** xii

# CHAPTER 1

# Research Background

1.1. General Introduction 1

1.2. What is attention? 2

1.3. How is attentional bias measured? 2

1.3.1. Emotional Stroop tasks 3

1.3.2. Visual probe tasks 3

1.3.3. Other tasks 4

1.4. What evidence links abnormalities of emotional attention and the emotional disorders? 5

1.4.1. Anxiety and depression 5

1.4.2. The causal role of attentional bias in the emotional disorders 8

1.4.3. Other disorders 8

1.4.4. Summary 9

1.5. What neural processes control the deployment of attention to emotional information? 9

1.5.1. Amygdala 10

1.5.2. Frontal control 10

1.5.3. Summary 13
1.6. Is the function of attentional control mechanisms altered in the emotional disorders? 14
1.7. Can the effects of treatment on cognition be disentangled from the effects of mood? 15
1.8. Pharmacological Methods of Modulating Attention to Emotional Information 16
1.8.1. Tryptophan depletion 16
1.8.2. Antidepressant administration 18
1.8.3. Alternative pharmacological strategies 20
1.8.4. Summary 21
1.9. Psychological Methods of Modulating Attention to Emotional Information 25
1.9.1. Cognitive Behavioural Therapy 25
1.9.2. Attentional bias training 25
1.9.3. Summary 27
1.10. Synthesis 33
1.10.1. Outstanding questions 33
1.10.2. Implications for treatment 34
1.10.3. Conclusion 35

# Chapter 2

# Characterising the Generalisation of Attentional Bias Training

2.1. Introduction 37
2.2. General Methodology 39
2.2.1. General Overview 39
2.2.2. Participants 41
2.3. Study 1 41
2.3.1. Methods 41
2.3.2. Results 46
2.3.3. Summary 53
2.4. Study 2 53
2.4.1. Methods 53

2.4.2. Results 54
2.4.3. Summary 56
2.5. Discussion 56
2.5.1. The basic attentional bias training effect 57
2.5.2. The generalisation of attentional bias training effects 57
2.5.3. Limitations 58
2.5.4. Conclusion 60

# Chapter 3

# Can Alteration of Prefrontal Function Account for the Effects of Attentional Bias training? An fMRI Study

3.1. Introduction 61
3.2. Methods and Materials 63
3.2.1. Participants 63
3.2.2. Questionnaire Measures 64
3.2.3. Attentional Bias training Task 64
3.2.4. Imaging Task 64
3.2.5. Image acquisition 66
3.2.6. Data Analysis 66
3.2.7. Connectivity analysis 68
3.3. Results 72
3.3.1. Questionnaire measures 72
3.3.2. Whole brain analysis 73
3.3.3. Connectivity Analysis 75
3.3.4. Region of interest analysis 76
3.3.5. Behavioural analysis 77
3.4. Discussion 77
3.4.1. Does the lPFC mediate the attentional bias training effect? 78
3.4.2. Involvement of other regions in the training effect 79

3.4.3. Is the prefrontal cortex responding to conflict or to the violation of an expectation? 80
3.4.4. Conclusion 81

# Chapter Four

# Threat Expectation and Surprise Predict Behavioural and Neural Measures of Attention to Emotional Stimuli

4.1. Introduction 82
4.2. Methods 84
4.2.1. Overview 84
4.2.2. Imaging Task 85
4.2.3. Computational Model 86
4.2.4. Data analysis 89
4.3. Results 92
4.3.1. Behavioural Data 92
4.3.2. Imaging data 93
4.4. Discussion 98
4.4.1. The generation of threat expectation in the brain 98
4.4.2. Alternative models 99
4.4.3. Does negative affectivity modulate attentional control? 100
4.4.4. Limitations to the interpretation of the model 101
4.4.5. Modelling the effects of attentional training 102
4.4.6. Conclusion 103

# Chapter Five

# Exploring the Interaction between Pharmacological and Cognitive Strategies for Altering Attentional Bias

5.1. Introduction 104
5.2. Methods 109

5.2.1. Summary of methodological changes in the current study 109
5.2.2. Overview 109
5.2.3. Participants 111
5.2.4. Questionnaire measures 112
5.2.5. Experimental Interventions 112
5.2.6. Assessment tasks 114
5.2.7. Statistical Analysis 117
5.3. Results 118
5.3.1. Demographic, compliance, side effect and baseline data 118
5.3.2. Did the interventions alter attentional function? 119
5.3.3. Did the interventions alter other measures of cognitive bias? 121
5.3.4. Did the interventions alter baseline measures of mood or anxiety? 127
5.3.5. Did the interventions alter emotional reactivity to external stressors? 127
5.3.6. Are the cognitive effects of the interventions related to their effects on emotional reactivity? 130
5.4. Discussion 131
5.4.1. The effects of the interventions on emotional memory and interpretation 131
5.4.2. The effects of the interventions on emotional reactivity to external challenge 134
5.4.3. Are the cognitive and emotional findings linked? 135
5.4.4. Limitations 136
5.4.5. Alternative treatment combinations 137
5.4.6. Conclusion 137

# Chapter 6

# Attentional Bias and its Modification in Patients with Bipolar Disorder

6.1. Introduction 138
6.2. Methods 141
6.2.1. Overview of study 141
6.2.2. Participants 142

6.2.3. Questionnaire assessment 143
6.2.4. Attentional bias training 144
6.2.5. Assessment of attentional bias 146
6.2.6. Emotional challenge task 146
6.2.7. Statistical analysis 147
6.3. Results 148
6.3.1. Demographic data and baseline anxiety 148
6.3.2. Patient characteristics 149
6.3.3. Baseline attentional biases 154
6.3.4. Effect of training on attentional bias within the patient group 155
6.3.5. Effect of training on emotional reactivity within the patient group 156
6.3.6. Effects of medication 156
6.4. Discussion 159
6.4.1. Baseline attentional bias 159
6.4.2. Effects of attentional training 162
6.4.3. Attentional control in bipolar disorder 163
6.4.4. Limitations 163
6.4.5. Conclusion 164

# Chapter 7

# General Discussion

7.1. Summary of findings 165
7.2. The mechanisms of action of attentional training 170
7.2.1. The level of the attentional training effect 171
7.2.2. Do psychological and pharmacological manipulations of attentional function target distinct control processes? 173
7.2.3. The nature of attentional biases in the emotional disorders 174
7.3. Clinical Implications 176
7.3.1. Attentional bias training as a treatment 176
7.3.2. The challenge of translational studies in clinical populations 178

7.4. Limitations 178
7.4.1. The assessment of attentional function 178
7.4.2. The use of healthy volunteers in treatment studies 179
7.4.3. Inferring causality using associational designs 180
7.5. Future Directions 181
7.6. Conclusion 182
**References 183**

# Acknowledgements

I am grateful to the Wellcome Trust for supporting me financially throughout the course of these studies. Particular thanks also go, obviously, to my three supervisors Catherine Harmer, Emily Holmes and Guy Goodwin who have been required, on a number of different occasions, to encourage, correct and tolerate me. Numerous conversations with Susie Murphy on the nature of attention, the brain and that sort of thing have been invaluable in developing the ideas described in this thesis. I would like to acknowledge the hard work of Maud Grol and Verena Ly in collecting the data reported in Chapter Five and the kindness of Tim Behrens for letting me use his computational model. More generally, this thesis could not have been completed without the input of Anjunabeats, Eurocamp, Wikipedia, Linux and Roast Chicken. Finally, I am, as ever, indebted to Nikki, Tom and Ben for things like purpose, meaning and fun.

# Declaration

Some of the work described in this thesis has previously been published.

**Chapter One**

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(1), 8-20.

**Chapters Two and Three**

Browning, M., Holmes, E. A., Murphy, S. E., Goodwin, G. M., & Harmer, C. J. (2010). Lateral prefrontal cortex mediates the cognitive modification of attentional bias. Biol Psychiatry, 67(10), 919-925.

# Abbreviations and Glossary

| | |
|---|---|
| ACC | Anterior Cingulate Cortex |
| ANOVA | Analysis of Variance |
| BDI | Beck Depression Inventory |
| BIS/BAS | Behavioural Inhibition Scale/Behavioural Activation Scale |
| BOLD | Blood Oxygenation Level Dependent; an MRI based signal which is used to infer local neural activation during fMRI |
| CBMa | Cognitive Bias Modification of Attention; the targeted alteration of habitual attentional deployment, referred to as attentional bias training in this thesis |
| CBMi | Cognitive Bias Modification of Interpretation; similar to CBMa although it aims to alter biases in interpretation rather than attention |
| CBT | Cognitive Behavioural Therapy |
| dACC | Dorsal Anterior Cingulate Cortex |
| EPI | Echo Planar Imaging; a form of MRI data acquisition often used in fMRI |
| ERP | Event Related Potential |
| FEAT | fMRI Experts Analysis Tool; an fMRI image analysis programme provided as part of the Functional Magnetic Resonance Imaging of the Brain (FMRIB)'s Software Library |
| FFA | Fusiform Face Area |
| fMRI | Functional Magnetic Resonance Imaging |
| lPFC | Lateral Prefrontal Cortex |
| MNI | Montreal Neurological Institute |
| MPRAGE | A MRI pulse sequence often used for the rapid acquisition of structural images |
| NHS | National Health Service |

| | |
|---|---|
| NICE | National Institute for Health and Clinical Excellence |
| OFC | Orbitofrontal Cortex |
| PPI | Psychophysiological Interaction; a form of fMRI analysis which examines the changes in functional connectivity between brain areas which are associated with altering task demands |
| PTSD | Post Traumatic Stress Disorder |
| rACC | Rostral Anterior Cingulate Cortex |
| ROI | Region of Interest |
| SD | Standard Deviation |
| SEM | Standard Error of the Mean |
| SSRI | Selective Serotonin Reuptake Inhibitors |
| STAI | Spielberger State-Trait Anxiety Inventory |
| TD | Tryptophan Depletion |
| TE | Echo Time; MRI acquisition parameter in spin echo sequences which corresponds to the time in ms between application of the 90° radiofrequency pulse and acquisition of the imaging data |
| TR | Repetition Time; MRI acquisition parameter corresponding to the time between successive pulse sequences being applied to the same slice |
| VAS | Visual Analogue Scale |
| Vigilance Score | A derived measure used in visual probe tasks as an estimate of attentional bias. |

# CHAPTER 1

# Research Background

## 1.1. General Introduction

A wealth of data from behavioural, neuroimaging, pharmacological and genetic studies has served to highlight the centrality of attentional processes in our understanding of anxiety and depression. Aberrant deployment of attention, particularly towards emotional information, occupies a pivotal position in many of the contemporary models of these disorders in which it is considered a causally relevant, proximal illness process. This view suggests that interventions which modify the habitual deployment of attention to emotional information should impact upon illness expression. Consistent with this proposal, both the pharmacological and psychological evidence based treatments of the emotional disorders have been shown to alter attentional function. More recently, the experimental modification of attention using "cognitive bias training" tasks[1] has been demonstrated to impact both mood and anxiety. These findings provide the foundation of the current thesis which has two broad aims; firstly, to investigate the mechanisms by which cognitive bias training tasks alter attention to emotional information and secondly, to use these initial results to improve our understanding of how the brain controls attention and to guide the development of novel

[1] These are sometimes referred to as "Cognitive Bias Modification of Attention" or CBMa tasks.

treatment strategies. This introductory chapter introduces the clinical and neuroanatomical models which seek to explain how attention to emotional information is controlled and then reviews the methods by which this control may be modified. The chapter concludes by highlighting the outstanding gaps in knowledge as well as summarising the implications of the literature, including an assessment of the opportunities for improving our aetiological understanding of the emotional disorders and developing novel treatment strategies. In the following chapters behavioural and neuroimaging studies that attempt to answer some of these outstanding questions and capitalise on the opportunities will be described.

## 1.2. What is attention?

Attention is the process by which cognitive resources are allocated preferentially towards or away from particular "objects or trains of thought" (James 1890). The fundamental nature of attention in cognition has resulted in numerous approaches and techniques being used to investigate it, the great majority of which are beyond the scope of this thesis. Instead, the focus of the work presented in the following chapters will be on the tendency for individuals to direct their attention towards, or away from certain types of emotional information; a habit that has been termed "attentional bias" (Mathews and MacLeod 2005). The great majority of work on attentional bias, including the studies in the current thesis, describe the covert (i.e. without requiring the redirection of the sense organs) deployment of attention to specific stimuli. Unless otherwise stated therefore, this is the sense in which the term attention will be used in the following chapters.

## 1.3. How is attentional bias measured?

A large range of behavioural tasks designed to measure the deployment of attention to emotional information have been described. Generally, target detection accuracy or reaction time are used in these tasks to infer attentional function as such measures are widely considered to be more reliable than the self-report of internal states (Nisbett and Wilson 1977). In the following section the rationale behind the most common of these tasks are summarised.

### 1.3.1. Emotional Stroop tasks

During standard emotional Stroop tasks (Figure 1.1A) participants are required to ignore the emotional content of a word stimulus while reporting its font colour, the rationale being that the emotional aspect of the stimulus interferes with the required behaviour and therefore slows response times. The demonstration of a specific slowing of response times to emotional stimuli is argued to demonstrate that more processing resources (i.e. attention) have been captured by the emotional dimension of the stimuli (Williams, Mathews et al. 1996) although this interpretation has been disputed (Algom, Chajut et al. 2004; Larsen, Mercer et al. 2006). Emotional Stroop tasks are insensitive to the timing of attentional effects, instead they provide a general indication that task interference has occurred at some point. A possible exception to this is when masked stimuli are presented; in this case the relevant stimulus is briefly presented and followed by a non-emotional mask which prevents explicit identification of the stimulus. In such cases it seems reasonable to assume that early stimulus appraisal processes lead to the task interference.

### 1.3.2. Visual probe tasks

Controversy in the interpretation of Stroop data has led to the increased popularity of visual probe tasks (Figure 1.1B). Based on paradigms developed by Posner and colleagues (Posner and Petersen 1990) these have been argued to provide a purer estimate of spatial attention. The most common variety of this task involves the brief presentation of emotional and neutral stimuli followed by a probe, to which the participant must respond, in the location of one of the stimuli (MacLeod, Mathews et al. 1986). It is assumed that reaction time is improved if attention is focused at the location in which the probe appears, that is if the stimulus previously found at that location had drawn attention. An estimate of attentional bias for emotional stimuli is calculated by comparing reaction times when the probe is found in the location of emotional stimuli compared to when it is found elsewhere. An advantage of visual probe tasks is that they provide information on the temporal characteristics of attention, specifically they have been described as providing an estimate of attentional deployment at the time the probe is presented (Cooper and Langton 2006). Recently, an alternative version of the visual probe task has been adapted to study visual attention (Fox, Russo et al. 2001). In this task a single emotional or neutral stimulus is used, with the probe appearing either in the same location as the stimulus or in an "uncued" (Posner 1980) location. It has been suggested that this paradigm is able to discriminate between the initial orientating and later disengagement phases of attention (Posner and Petersen 1990; Fox, Russo et al. 2001; Koster, Crombez et al. 2007; Baert, De Raedt et al. 2010).

### 1.3.3. Other tasks

Although the above tasks are the most commonly used in the attentional bias literature, a large number of other measures of attention have been described. These include tasks which

assess attention in other modalities than vision, such as the dichotic listening task (Sander, Grandjean et al. 2005), tasks which measure the temporal rather than spatial deployment of attention such as the attentional blink task (Anderson 2005; Fox, Russo et al. 2005; De Martino, Strange et al. 2008) and tasks which have been used to differentiate the early, parallel from later, serial phases of attentional filtering, such as visual search (Hansen and Hansen 1988; Dandeneau, Baldwin et al. 2007). Finally the overt deployment of attention, during which the relevant sensory organs are redirected, can also be measured directly using techniques such as infrared eye tracking (Derakshan, Salt et al. 2009).

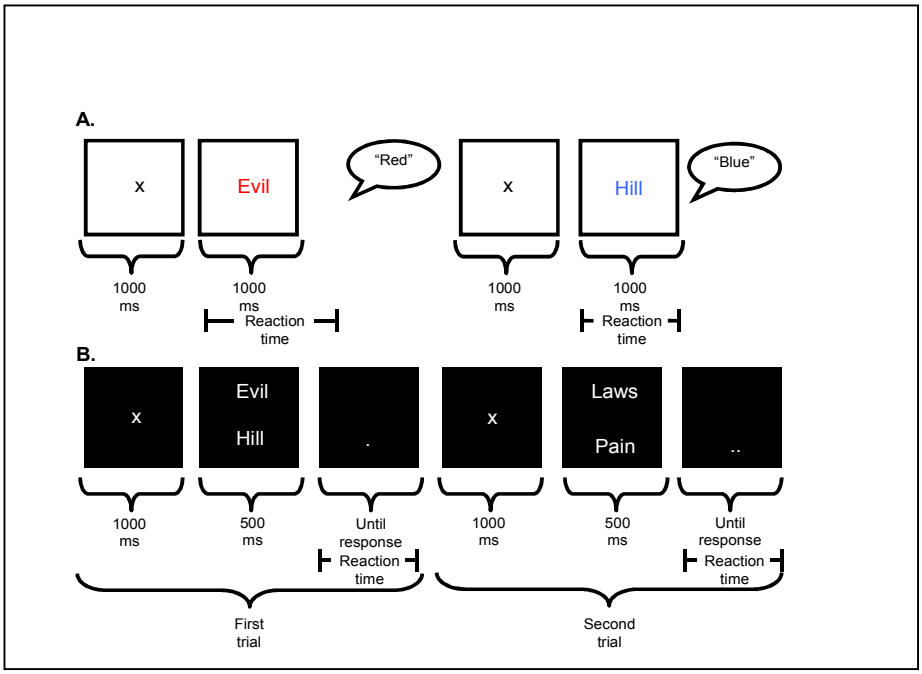



**Figure 1.1. Behavioural methods of assessing emotional attention. Example trials from the two most common behavioural paradigms used to assess attentional function; the emotional Stroop (A) and visual probe (B) tasks.**

## 1.4. What evidence links abnormalities of emotional attention and the emotional disorders?

### 1.4.1. Anxiety and depression

Both depression and anxiety have been associated with biases in the processing of emotional information; patients with these disorders habitually interpret, attend to and/or remember information in a more negative manner than non-clinical control participants (Mathews and MacLeod 2005). Cognitive theorists suggest that these habits of thought are causal factors in the aetiology and maintenance of the disorders (Beck 1976; Williams, Watts et al. 1997; Mathews and Mackintosh 1998; Mogg and Bradley 1998; Mathews and MacLeod 2005). Anxiety has particularly been associated with negative attentional biases (see Section 1.2 and Figure 1.1 for a summary of the common tasks used to assess attentional bias), with a number of studies demonstrating an association between anxiety and negative attentional biases in both clinical and non-clinical groups across a range of stimuli and measures of attention (Mathews and MacLeod 1985; Williams, Mathews et al. 1996; Fox, Russo et al. 2001; MacLeod, Campbell et al. 2004; Fox, Russo et al. 2005).

In depression the most consistent cognitive abnormalities were initially described in measures of memory (Williams, Watts et al. 1997; Gilboa-Schechtman, Erhard-Weiss et al. 2002; Ridout, Astell et al. 2003; Mathews and MacLeod 2005), however there has been increasing interest in the possibility that attentional biases may exist in this disorder as well. Specifically, depression has been shown to be associated with a tendency to attend to the spatial location of negative words (Mogg, Bradley et al. 1995; Bradley, Mogg et al. 1997; Donaldson, Lam et al. 2007) and faces (Gotlib, Kasch et al. 2004; Gotlib, Krasnoperova et al. 2004; Joormann and Gotlib 2007; Leyman, De Raedt et al. 2007). Negative attentional biases are also demonstrable in previously depressed, currently euthymic patients (Joormann and

Gotlib 2007) lending support to the proposal that they are involved in the vulnerability to depression rather than solely reflecting a state marker or symptom of the illness.

Although depression and anxiety have both been associated with attentional bias towards negative stimuli there are some differences in how this bias manifests between the disorders. Most obviously there appears to be a variation in how long a stimulus must be presented before the bias becomes evident; whereas anxiety is associated with an attentional bias occurring from 10 to 500 ms after stimulus presentation, in depression the bias seems to occur at around 1000 ms (note that these timings refer to the duration of stimulus presentation rather than to neuronal latencies; Gotlib, Kasch et al. 2004). These timing effects are consistent with dual process accounts of brain organization which describe the existence of two distinct information processing streams; an automatic stream which processes information in a rapid but inflexible manner and a strategic stream which provides slower, more flexible processing (see Carver, Johnson et al. 2008 for a recent review of the dual process models relevant to the emotional disorders). It has been suggested (Mathews and MacLeod 2005) that anxiety is associated with abnormalities in the early, automatic processing of information and depression with abnormalities in later processing stages. Despite the differences seen in experimental paradigms, it is notable that individuals very often experience both anxious and depressive symptoms concurrently (Maser and Cloninger 1990; Mineka, Watson et al. 1998) and that there is a large overlap in the evidence based treatments for these disorders (NICE 2007; NICE 2009). Thus while there are some differences in the expression of attentional biases in anxiety and depression there remains a strong possibility that these disorders share at least some pathological processes.

### 1.4.2. The causal role of attentional bias in the emotional disorders

As the studies described above are uniformly observational they provide relatively weak evidence that negative attentional biases are causally related to the emotional disorders. However, alternative study designs which are more able to directly address questions of causality have also been completed. Firstly, in a longitudinal study (MacLeod and Hagan 1992) it was demonstrated that attentional bias towards threat as measured by an emotional Stroop task predicted anxiety symptoms provoked by a subsequent naturalistic stressor (receiving worrying results following cervical colposcopy) suggesting that the negative attentional biases temporally precede the associated anxiety symptoms. Most convincing however is a recent body of work, described in detail later in this chapter (section 1.9.2), which has employed simple computer based tasks in order to experimentally induce attentional biases in non-clinical populations. These experimental studies provide more direct support for a causal role of attention in the emotional disorders by demonstrating that inducing a negative attentional bias can lead to symptoms of anxiety and depression in non-clinical participants.

### 1.4.3. Other disorders

Attentional biases are not limited to the emotional disorders; across a range of other diagnoses attentional biases have been described to stimuli which are salient for the particular population being studied, such as painful stimuli in chronic pain (Van Damme, Legrain et al. 2010) or food related stimuli in the eating disorders (Lee and Shafran 2004). Similarly, attentional biases towards negative emotional material are not limited to individuals with diagnoses of anxiety or depression; rather they increase as a function of anxious and

depressive symptoms, even in non-clinical participants (Bradley, Mogg et al. 1997; Fox, Russo et al. 2001; Koster, Crombez et al. 2006). This observation suggests that negative attentional biases may be related to mood and anxiety symptoms regardless of diagnoses (Harvey, Watkins et al. 2004). If true this would suggest that the symptoms of anxiety and depression found in other axis I disorders, such as Bipolar Disorder (Simon, Otto et al. 2004; McIntyre and Keck 2006), are also related to attentional function. However, empirical evidence for a causal relationship between emotional symptoms and attentional bias in such disorders is currently lacking.

### 1.4.4. Summary

In summary, there is evidence that an increased tendency to direct attention towards negative information leads to increased symptoms of depression and anxiety. Indirect evidence suggests that the emotional symptoms found in other patient groups may also arise as a consequence of attentional biases although direct evidence for this is currently lacking. Differences in the timing of the attentional biases found in depression and anxiety can be accounted for by dual process models which postulate the existence of temporally separable attention control processes.

## 1.5. What neural processes control the deployment of attention to emotional information?

The control of attention to emotional information is often conceptualised within a biased competition framework (Desimone and Duncan 1995; Mathews and Mackintosh 1998; Pessoa and Ungerleider 2005; Vuilleumier 2005; Bishop 2007). Broadly, this suggests that

neural representations of stimuli compete for processing resources; success in securing these resources is seen as a neural marker of attention to the stimulus. Support for this view comes from both single cell recording in non-human primates and neuroimaging studies in human populations in which the sensory cortical response to a stimulus is increased when it is attended (Luck, Chelazzi et al. 1997; Hopfinger, Buonocore et al. 2000; Driver 2001). The competition for processing resources is influenced by biasing signals which act to favour or penalise specific representations and thus to encourage the focus of attention towards or away from particular stimuli. Two neural systems have been implicated in generating signals of particular relevance to emotional stimuli.

### 1.5.1. Amygdala

An amygdala based system is thought to produce a fast, relatively automatic signal which promotes the deployment of attention towards salient stimuli (Adolphs, Tranel et al. 1995; Whalen, Rauch et al. 1998; Pessoa, Padmala et al. 2005; Lim, Padmala et al. 2009). While some authors suggest that the amygdala response to emotional stimuli is fundamentally automatic, occurring regardless of the perceptual or cognitive load in which it is presented (Öhman 2002), others suggest that, in line with recent accounts of attentional selection (Lavie 2005), the amygdala is not immune to these considerations (Pessoa, Padmala et al. 2005; Bishop, Jenkins et al. 2007; Lim, Padmala et al. 2009). However, regardless of whether amygdala activity is strictly automatic, there is general agreement that it provides a fast signal which reflects stimuli salience.

### 1.5.2. Frontal control

A more flexible and strategic response is associated with a second signal, originating in areas of the prefrontal cortex including both the anterior cingulate cortex (ACC) and the lateral prefrontal cortex (lPFC). These regions are conceptualised as providing top-down control of cognition which allows attention to be maintained on relevant stimuli, even in the presence of distraction. Prefrontal control impacts a broad range of cognitive processes (e.g. Duncan and Owen 2000; Monsell 2003), including the deployment of attention to both emotional and non-emotional stimuli (MacDonald, Cohen et al. 2000; Bishop, Duncan et al. 2004a; Etkin, Egner et al. 2006). While there appears to be a degree of overlap in the systems involved in the control of attention to emotional and non-emotional stimuli, there is evidence of a functional distinction within the ACC. Specifically, it has been suggested (Bush, Luu et al. 2000) that the rostral portion of the ACC (rACC), which lies anterior to the genu of the corpus callosum is sensitive to emotional dimensions of sensory information, whereas the more dorsal portion (dACC) responds to non-emotional information. This distinction appears to hold for emotional attention, with the majority of studies reporting results specifically from the rACC (Bishop, Duncan et al. 2004a; Etkin, Egner et al. 2006).

*1.5.2.1. Conflict*

One of the most influential attempts to delineate the specific functions of the ACC and lPFC suggests that these areas are concerned with the identification of and response to "conflict" (Botvinick, Cohen et al. 2004). Conflict was initially defined as arising when the representations of incompatible *responses* were simultaneously activated (Botvinick, Braver et al. 2001). It was suggested that the ACC acts primarily to detect conflict which triggers the recruitment of control regions such as the lPFC in order to constrain processing and thus to

reduce subsequent conflict (MacDonald, Cohen et al. 2000). However, it has proven difficult to account for all experimental observations using this initial formulation (Jones, Cho et al. 2002; Nachev, Rees et al. 2005; Botvinick 2007; Egner 2007); for example, both behavioural interference and ACC activity in response to a distractor is greater if the distractor occurs infrequently (Jones, Cho et al. 2002; Bishop, Duncan et al. 2004a). Conversely, frequent distractors result in decreased interference, and when sufficiently common may even be associated with less interference than trials in which they are absent (Logan and Zbrodoff 1979; Sturmer, Leuthold et al. 2002) a finding not easily explained by the conflict model. This observation has lead some authors to suggest that the conflict construct should be expanded to take into account recent stimulus history (Jones, Cho et al. 2002), however alternative accounts of PFC function may provide a more natural explanation for such results (Egner 2007).

*1.5.2.2. Expectation*

An alternative model of cognitive control describes the process as generating and tracking "expectancies" (Logan and Zbrodoff 1979; Egner 2007); for example, with regard to the stimulus frequency effect described above, participants generate an expectation of the stimuli which are likely to occur and adapt their processing accordingly, the increased interference provoked by an infrequent distractor arises because, in that context, the distractor is unexpected. A number of recent accounts incorporate insights from both the associative (Rescorla and Wagner 1972) and machine learning literatures (Sutton and Barto 1998), to highlight the role of the PFC in generating and tracking such expectancies (Rushworth, Mars et al. 2009). These models suggest that expectations (sometimes called predictions) are

updated using a derived parameter, the prediction error, which is computed as the difference between current predictions and reality. The literature is most developed with regard to reward prediction with evidence that the values of actions and stimuli (the expected reward associated with actions or stimuli) are encoded in medial and orbital frontal cortex (Rudebeck, Behrens et al. 2008; Walton, Behrens et al. 2010) and that the reward prediction error is encoded in the phasic firing of dopaminergic neurons originating in the ventral tegmental area (Schultz, Dayan et al. 1997). However, it has been proposed that the updating of predictions by prediction errors may provide a more general account of cortical organisation (Friston, Kilner et al. 2006) and thus explain a greater range of behaviour and brain activity than just that associated with reward. Supporting this assertion, recent evidence suggests that such prediction error signals occur in unrewarded, stimulus-stimulus associative learning (den Ouden, Friston et al. 2009) and that both prediction and prediction error signals are apparent in the visual cortex during perceptual tasks (Summerfield and Koechlin 2008). Although no neuroimaging studies have specifically investigated expectancy mechanisms in attentional tasks a behavioural study indicates that attention is deployed preferentially towards visually surprising stimuli (i.e. those with a large prediction error; Itti and Baldi 2009).

### 1.5.3. Summary

In summary, activity of sensory and related association cortices provides a neural marker of the deployment of attention to a stimulus. Attention to emotional stimuli is particularly influenced by an amygdala based signal which reflects the salience of the stimuli and by a frontal signal which facilitates a more flexible deployment of attention. The specific roles of

the various frontal regions involved in attentional control is disputed although both conflict and expectancy based models have empirical support.

## 1.6. Is the function of attentional control mechanisms altered in the emotional disorders?

The above neuroanatomical model predicts that the negative attentional biases characteristic of emotional disorders may arise from perturbation of either of the amygdala or frontal control systems. Supporting this contention, experimental studies report increased amygdala activity to negative stimuli in both depression (Sheline, Barch et al. 2001; Siegle, Steinhauer et al. 2002; Fu, Williams et al. 2004; Fales, Barch et al. 2008) and anxiety (Schneider, Weiss et al. 1999; Tillfors, Furmark et al. 2002; Bishop, Duncan et al. 2004b; Shin, Rauch et al. 2006; Straube, Mentzel et al. 2006) whereas frontal activation to the same stimuli has been found to be decreased in both disorders (Mayberg, Liotti et al. 1999; Shin, Whalen et al. 2001; Tillfors, Furmark et al. 2002; Bishop, Duncan et al. 2004a; although see also Straube, Mentzel et al. 2006; Fales, Barch et al. 2008). A notable discrepancy arises when comparing this data to the behavioural literature; whereas differences in the temporal characteristics of the biases associated with anxiety and depression are evident when assessed behaviourally (see section 1.4.1), neuroimaging studies do not seem to reliably differentiate the disorders. This may be due to constraints on the temporal resolution of functional imaging techniques or the fact that neuroimaging studies have tended to simply present emotional stimuli without manipulating participants' attention (notable exceptions to this rule being; Bishop, Duncan et al. 2004a; Etkin, Egner et al. 2006). Such manipulations allow the effects of attention to be isolated from those due simply to the emotional properties of a stimulus. Nevertheless, the

temporal specificity of the attentional effect demonstrated in the behavioural studies of anxiety versus depression remains unexplained by current neuroimaging data.

Generally however, the neuroimaging data compliment the behavioural findings in that the most commonly reported neuroimaging abnormalities in the emotional disorders would be expected to lead to negative attentional biases. These abnormalities, increased amygdala and decreased frontal activation in response to attended negative stimuli, are evident in both depression and anxiety. This leads to the prediction that interventions which alter the function of either one or both of these systems will influence emotional attention and, in turn, current anxious and depressive symptoms.

## 1.7. Can the effects of treatment on cognition be disentangled from the effects of mood?

The interpretation of studies which investigate the effects of treatments on cognition can, paradoxically, be hampered when the treatment is effective. The problem is particularly acute in studies of clinical populations in which treatments are expected to improve clinical state (e.g. Mathews, Mogg et al. 1995; Kennedy, Evans et al. 2001; Fu, Williams et al. 2004; Fales, Barch et al. 2009). These studies inevitably compare groups which differ on two accounts; exposure to treatment (e.g. administration of antidepressant vs. placebo) and current levels of psychopathology (e.g. mood). While such studies are essential to examine the effects of a treatment on the relevant population, it is difficult to be certain whether any differences observed in behaviour or neural activity reflect the direct action of the intervention or are general effects of clinical status. The influence of this confounding factor

can be reduced by studying non-clinical populations who generally do not experience profound changes of mood and who can thus provide important complementary data to the clinical studies. As the critical issue in the following sections concerns the mechanisms by which interventions may alter attentional function both studies of clinical populations and those solely involving non-clinical participants will be reviewed.

## 1.8. Pharmacological Methods of Modulating Attention to Emotional Information

Neurochemical probes have been used to examine how neurotransmitter function may modulate attentional bias both in healthy volunteers and patients. One key strategy has been to see whether biases apparent in depression and anxiety can be mimicked by depletion of the neurochemicals known to be involved in these disorders.

### 1.8.1. Tryptophan depletion

The synthesis of serotonin in the brain is dependent on the availability of its precursor amino acid, tryptophan, from plasma. Acute administration of an amino acid mixture that selectively lacks tryptophan is effective in decreasing availability of tryptophan to the brain through processes of increased protein synthesis (lowering plasma tryptophan levels) and increased competition for transport across the blood-brain barrier (Reilly, McTavish et al. 1997). Tryptophan depletion (TD) transiently lowers mood in recovered depressed patients, although there is no consistent effect on mood in healthy volunteers (Ruhe, Mason et al. 2007). Recently, however, TD has been used as an experimental tool to examine the role of serotonin in the processing of emotional information in healthy volunteers (Booij, Van der

Does et al. 2003). That TD is seen to affect cognition in non-clinical participants without altering mood suggests that cognitive measures, such as the attentional tasks examined in this chapter, provide a more proximal or, at least, more sensitive marker of the effects of this intervention than mood measures. A number of studies (see Table 1.1) have suggested that reduction of serotonergic function using TD increases the interference associated with negative words in emotional Stroop tasks, both in never (Evers, van derVeen et al. 2006) and previously depressed populations (Hayward, Goodwin et al. 2005; Munafò, Hayward et al. 2006). Interestingly, a study in which participants took tryptophan supplementation, an intervention predicted to increase serotonergic function, suggested that attention towards negative words was decreased in females (Murphy, Longhitano et al. 2006). However, other studies have produced inconsistent results. First, in previously depressed participants TD was found to slow responses in the emotional Stroop task to positive words, suggesting increased attentional interference from these positive cues (Booij, Van der Does et al. 2005). A further study found no effect of TD on a visual-probe task involving threat related words (Merens, Booij et al. 2008). Thus, while TD demonstrably influences the processing of emotional information, there is a degree of inconsistency to this effect.

Two neuroimaging studies have investigated the effects of TD on tasks in which emotional valence was manipulated, although neither explicitly manipulated attention. Both reported increased amygdala activation in response to negative stimuli following TD, although this effect was qualified by participant characteristics, which may explain some of the variance in the behavioural data. Thus in a non-clinical sample TD increased amygdala activity to fearful vs. happy faces but only in those with high scores on the behavioural inhibition

subscale of the BIS/BAS measure (Carver and White 1994; Cools, Calder et al. 2005). Similarly, a second study reported increased amygdala activation following TD in a group of healthy women, but only those with a family history of depression (van der Veen, Evers et al. 2007).

Taken as a whole, these findings suggest that TD increases attention to negative emotional information and that its effect is associated with alteration of amygdala function; however, as has been noted using non-attentional measures of cognition (Robinson, Cools et al. 2009), this effect is qualified by individual differences in the participants.

### 1.8.2. Antidepressant administration

An alternative strategy in the investigation of the effects of neurotransmitter function on attentional bias is to characterise the effects of the pharmacological treatments of the emotional disorders. Serotonergic function can be potentiated using antidepressant medications which specifically block the reuptake of serotonin from the synapse, a strategy recommended in the treatment of both depression and anxiety (Nutt 2002; NICE 2007; NICE 2009). A number of experimental studies have assessed the influence of these serotonergic antidepressants on emotional attention in non-clinical groups (see Table 1.1). A single dose of the antidepressant citalopram was associated with an increased bias towards positive words in a visual-probe task (Browning, Reid et al. 2007), whereas seven days of citalopram decreased attention to briefly presented fearful faces, an effect not seen with the norepinephric antidepressant reboxetine (Murphy, Yiend et al. 2009). As predicted and contrary to the effects of TD, serotonergic antidepressants decrease amygdala activation to

negative stimuli in both clinical (Sheline, Barch et al. 2001; Fu, Williams et al. 2004) and non-clinical populations (Del-Ben, Deakin et al. 2005; Harmer, Mackay et al. 2006; Arce, Simmons et al. 2008; Murphy, Norbury et al. 2009). There is less evidence that antidepressants alter activity in frontal control regions, two studies report such an effect (Kennedy, Evans et al. 2001; Fales, Barch et al. 2009) although both of these involve clinical populations indicating that the findings may be secondary to improved clinical status rather than treatment effect. An event related potential (ERP) study has examined the effects of citalopram and reboxetine on the processing of emotional information in healthy volunteers (Kerestes, Labuschagne et al. 2009). The high temporal acuity of ERP allowed the authors to demonstrate that both antidepressants influenced processing relatively shortly after stimulus presentation (250ms) with the effect being found in a component of the ERP known to be sensitive to spatial attention (Eimer, Holmes et al. 2003). The neuroimaging findings are thus consistent with an initial effect of antidepressants on the amygdala, with alteration of frontal control regions occurring later on in treatment and potentially reflecting improvement in clinical status. Interestingly, a single study has demonstrated that reboxetine also reduces amygdala activity to negative stimuli (Norbury, Mackay et al. 2007) despite apparently having no effect on spatial attention to threat (Murphy, Yiend et al. 2009). A possible explanation of this discrepancy is provided by recent models of norepinephric function which suggest that it acts specifically as a temporal filter in the control of attention (Aston-Jones and Cohen 2005). Thus the behavioural effects of norepinephric manipulation may best be assessed using measures of the deployment of attention in time rather than space. De Martino and colleagues (De Martino, Strange et al. 2008) report a series of three studies which

examine the influence of the beta-blocker propranolol and reboxetine on an emotional version of the attentional blink task, which specifically assesses the deployment of attention in time. The authors report a generally deleterious effect of propranolol on attentional function in that it caused poorer task performance regardless of stimuli salience. Reboxetine was found to improve task performance, although only in trials involving negative emotional stimuli. No positive stimuli were used in the study limiting conclusions as to the specificity of reboxetine's effect although they would be in keeping with its effect on attention being primarily evident in the temporal domain.

Together these studies indicate that serotonergic antidepressants alter attention to emotional information by encouraging a relatively more positive bias. Again these effects are associated with an effect of the medication on amygdala function.

### 1.8.3. Alternative pharmacological strategies

Based on clinical findings suggesting that acute administration of glucocorticoids reduce the symptoms of Post Traumatic Stress Disorder (PTSD) (Schelling, Roozendaal et al. 2004), Putman and colleagues (Putman, Hermans et al. 2007) used an emotional Stroop task to demonstrate that a single dose of the glucocorticoid cortisol reduced interference from masked negative stimuli (fearful faces) in healthy men. The same group had previously reported a similar effect when a single dose of testosterone, which has antidepressant activity (Pope, Cohane et al. 2003), was administered to females (van Honk, Peper et al. 2005). Lastly a single dose of the benzodiazepine diazepam, which has acute anxiolytic effects, has been found to increase attentional bias towards masked happy faces in a non-clinical population

(Murphy, Downham et al. 2008) with a complementary effect of reduced amygdala activation to emotional stimuli being reported for another benzodiazepine, lorazepam (Paulus, Feinstein et al. 2005).

### 1.8.4. Summary

A number of pharmacological interventions have been shown to alter attention to emotional stimuli. Taken as a whole the behavioural data suggest that interventions which improve anxiety or depression (tryptophan supplementation, antidepressant medication, steroids, benzodiazepines) tend to result in attention being directed away from negative and towards positive stimuli. Tryptophan depletion, which leads to a worsening of symptoms in vulnerable individuals, tends to have the opposite effect on attention. As can be seen from Table 1.1, the majority of these effects are seen early on in the deployment of attention, with a number of studies demonstrating an effect with masked stimuli. The timing of the attentional effects suggests that these interventions are altering the function of the bottom-up, stimulus appraisal system; a conclusion supported by the neuroimaging studies in which altered amygdala activity in response to negative stimuli is the most common finding.

**Table 1.1: Studies Which Have Reported the Effects of Pharmacological Manipulations on Measures of Attention to Emotional Information.**

| Paper | Intervention | Population Studied [a] | Attentional Task | Type and Duration of Stimulus | Duration of Stimulus [b] | Result | Notes |
|---|---|---|---|---|---|---|---|
| (Evers, van derVeen et al. 2006) | Tryptophan depletion | Non-clinical females (n=19 using a within subject design) | Emotional Stroop | Positive, negative and neutral words | 200ms | Increased interference of negative words after TD | Result derived from error rate. No effect for reaction time |
| (Booij, Van der Does et al. 2005) | Tryptophan depletion | Patients in complete or partial remission from depression who were still taking SSRI medication (n=20 using a within subject design) | Emotional Stroop | Positive, negative and neutral words | Up to 1500mms | Increased interference from positive words | Complex experimental design with participants completing cognitive tests on four separate occasions |
| (Hayward, Goodwin et al. 2005) | Tryptophan depletion | Patients (n=24) with a history of depression who no longer met diagnostic criteria and healthy controls (n=24) | Emotional Stroop | Negative and neutral words | 400ms | Increased interference from negative words | Effect seen in both clinical and control group. Counting version of the emotional Stroop task used. |
| (Munafò, Hayward et al. 2006) | Tryptophan depletion | Patients with previous depression who were on antidepressant medication (n=24), no medication | Emotional Stroop | Negative and neutral words | Unmasked (until response) or **masked (14ms)** | Increased interference from negative words in both masked and | Effect only seen in previously depressed group who were still taking medication |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | (n=24) and non-clinical controls (n=24) | | | conditions | unmasked conditions | |
| (Merens, Booij et al. 2008) | Tryptophan depletion | Patients (n=16 using a within subject design) in complete or partial remission from depression who were still taking antidepressant medication | Visual probe | Neutral vs. threat words and depression relevant vs. positive words. | 500ms | No significant effects. | |
| (Murphy, Longhitano et al. 2006) | Tryptophan Supplementation | Non-clinical females (n=19 per group) | Visual probe | Positive, negative and neutral words | **Unmasked (500ms)** and masked (14ms) conditions | Decreased attention to negative words in unmasked condition | Result derived from error rate. No effect for males. |
| (Browning, Reid et al. 2007) | Single dose of citalopram or placebo | Non-clinical group (n= 15 per group) | Visual probe | Positive, negative and neutral words. | **Unmasked (500ms) and masked (14ms)** conditions | Increased attention to positive words in both masked and unmasked conditions | |
| (Murphy, Yiend et al. | Seven days of citalopram, reboxetine or | Non-clinical group (n=14 per group) | Visual probe | Happy, fearful or neutral facial | **Unmasked (100ms)** and masked (16ms) | Citalopram reduced attention to fearful face in unmasked | Effect seen only for citalopram, no effect of reboxetine |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 2009) | placebo | | | expressions | conditions | condition | |
| (Putman, Hermans et al. 2007) | Cortisol or placebo | Non-clinical males (n=20 using within subject design) | Emotional Stroop | Fearful or neutral facial expressions | **Masked (14ms)** presentation | Decreased interference from fearful faces | |
| (van Honk, Peper et al. 2005) | Testosterone or placebo | Non-clinical females (n=16 using a within subject design) | Emotional Stroop | Fearful, happy or neutral facial expression | **Masked (14ms)** presentation | Decreased interference from fearful faces | No effect on happy faces. No mention of the effect of trial order in relevant analysis |
| (Murphy, Downham et al. 2008) | Diazepam or placebo | Non-clinical volunteers (n=12 per group) | Visual Probe | Fearful, happy or neutral facial expressions | Unmasked (100ms) or **masked (14ms)** presentation | Increased attention towards happy in masked condition | Authors argue that increased attention towards happy is similar to decreased negative bias as in the relevant trials happy and neutral stimuli are presented together and the neutral stimuli is relatively more threatening. |

[a] Unless otherwise stated studies used a between subjects design.

[b] Certain task designs (see figure one) allow inference as to the timing of the temporal effects of an intervention on attention. In studies which employ such a design and report significant findings the relevant duration is presented in bold

## 1.9. Psychological Methods of Modulating Attention to Emotional Information

### 1.9.1. Cognitive Behavioural Therapy

Cognitive Behavioural Therapy (CBT) is a complex psychological intervention which has been developed specifically to change the patterns of thinking and behaviour which impact on how an individual interacts with the environment and their resulting emotional state (Hollon, Stewart et al. 2006). CBT is recommended as a first line treatment for both depression and anxiety (NICE 2007; NICE 2009). A single study examining the effect of CBT on attentional function in anxiety disorder suggests that it reduces the negative attentional bias found pre-treatment (Mathews, Mogg et al. 1995). However, as discussed above, such a finding may be attributed to non-specific effects of clinical status rather than as necessarily a direct effect of treatment. The same observation applies to the neuroimaging studies of psychological therapies all of which have recruited clinical groups. These studies report changes of activation primarily in frontal regions (Brody, Saxena et al. 2001; Paquette, Levesque et al. 2003; Goldapple, Segal et al. 2004; Straube, Mentzel et al. 2006; Schienle, Schafer et al. 2007; Fu, Williams et al. 2008), but also in the amygdala (Fu, Williams et al. 2008) and basal ganglia (Martin, Martin et al. 2001).

### 1.9.2. Attentional bias training

Unfortunately, the complexity and variability of CBT can limit the extent to which it may be applied in controlled experimental trials. Perhaps because CBT is difficult to administer in

experimental studies, there has been an increasing interest in developing simpler cognitive tasks which target the key hypothesized emotional processes and which lend themselves to controlled experimentation. While many of these tasks involve the explicit and effortful control of emotion (Ochsner and Gross 2005; Phan, Fitzgerald et al. 2005; McRae, Hughes et al. 2009), a number of techniques have been developed recently to specifically alter the habitual deployment of attention to emotional information (MacLeod, Rutherford et al. 2002; Dandeneau, Baldwin et al. 2007; Baert, De Raedt et al. 2010). The most commonly used technique (MacLeod, Rutherford et al. 2002) employs a variant of the visual probe task (see Figure 1.1.1) in which the location of the probe is constrained such that it always replaces the neutral (for "avoid-threat" training) or negative (for "attend-threat" training) word. The task involves many experimental trials over the course of which participants learn to direct attention towards the type of stimuli which predict the probe location; e.g. if the probe always replaces negative stimuli participants develop the habit of attending to negative stimuli generally, a negative attentional bias.

Importantly there is increasing evidence that using such tasks to encourage a positive attentional bias ("avoid-threat" training) in clinical populations results in improvement in symptoms (Dandeneau, Baldwin et al. 2007; Eldara, Ricona et al. 2008; Hazen, Vasey et al. 2008; Amir, Beard et al. 2009; Schmidt, Richey et al. 2009). A number of studies have assessed the effects of these so called "attentional bias training" tasks on measures of emotional attention (see Table 1.2) in both clinical and non-clinical populations. Generally the tasks are seen to influence attention in the manner expected although, as can be seen from Table 1.2, they have only ever been shown to influence attention to unmasked stimuli

presented for 500 ms or longer. A caveat to this observation is that the training tasks themselves generally involve a version of the visual probe task in which stimuli are presented for 500 ms, raising the possibility that earlier effects on attention may be observed if the training task was altered to present stimuli for a shorter time. However, the published data to date differ from those involving pharmacological manipulations in which effects are reported during earlier stages of information processing. If this behavioural observation is valid it would suggest that cognitive and pharmacological interventions may alter distinct mechanisms of attentional control, a hypothesis that has previously been raised by a number of authors (DeRubeis, Siegle et al. 2008; Harmer 2008). Assessing the effects of attentional bias training on neural activity would directly test this hypothesis, although to date, no such studies have been performed.

### 1.9.3. Summary

Cognitive training tasks have been shown to alter both attention to emotional information and clinical status in a predictable manner, though further studies deploying the experimental tasks in clinical populations are needed. The extant results suggest that this effect occurs later in the deployment of attention than that found with pharmacological manipulations raising the possibility that different attentional control mechanisms are targeted by psychological and pharmacological interventions.

**Table 1.2: Studies which have reported the effects of psychological manipulations on measures of attention to emotional information.**

| Paper | Intervention | Population studied [a] | Attentional task | Type of stimulus | Duration of stimulus [b] | Result | Notes |
|---|---|---|---|---|---|---|---|
| (Mathews, Mogg et al. 1995) | Seven session group CBT for anxiety management | Patients with GAD ( n=24) were compared to non-clinical controls (n=23) both before and after treatment | Emotional Stroop and visual search | Negative and neutral words | Until response | Borderline significant reduction of negative attentional bias in the anxious group using emotional Stroop (p=0.06) and visual search (p=0.07) | Emotional Stroop task used cards rather than computer to present stimuli. |
| (MacLeod, Rutherford et al. 2002) | Single session of attend-threat or avoid-threat training. | Non-clinical (n=32 per group) | Visual probe | Negative and neutral words | **Unmasked (500ms)** and masked (20ms) | Training influenced attention in unmasked condition | Two studies performed, both demonstrating similar effect. |
| (Hazen, Vasey et al. 2008) | Five sessions of avoid-threat or sham attentional bias training | Non-clinical (n=12 per group) sample with high score on worry scale (includes participants who would meet diagnostic criteria for GAD) | Visual Probe | Negative and neutral words | **500ms** | Avoid-threat training reduced attention to negative words | |
| (Eldara, Ricona et al. 2008) | Single session of attend-threat or avoid-threat attentional bias | Non-clinical (n=13 per group) children aged 7-12 | Visual probe | Angry and neutral faces | **700ms** | Attend-threat training influenced attention | This effect was only significant for the faces used in training itself—no generalization to novel faces. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | training. | | | | | | |
| (Wadlinger and Isaacowitz 2008) | Single session of attend-positive or avoid-positive training | Non-clinical volunteers (approx n=22 per group) | Eye tracking | Positive and neutral words | 8000ms | Positive training caused decreased fixation on negative areas of images | Attentional bias training was visual probe type |
| (Li, Tan et al. 2008) | 7 sessions of daily avoid-threat/attend-happy or sham training | Non-clinical volunteers with high score on a social anxiety scale (n=12 per group) | Visual probe | Threatening and happy faces | **500ms** | Active training caused significant decrease in negative bias | Significant differences in measures of anxiety between groups at end of training may confound interpretation of attentional effects. The negative pictures were selected as being "threatening" rather than displaying a specific emotion. As the threatening pictures were paired with positive pictures these results could also be interpreted as showing an increase in positive attentional bias. |
| (See, Macleod et al. 2009) | 14 sessions of daily avoid-threat or sham training | Non-clinical volunteers (approx n=19 per group) | Visual probe | Negative and neutral words | **500ms** | Avoid-threat training caused significant decrease in negative bias | Significant differences in trait and state anxiety between groups at end of training may confound interpretation of attentional effects |
| (Amir, Beard et al. 2009) | 8 sessions of avoid-threat or sham training over 4 weeks | Patients with GAD (approx n=14 per group) | Visual probe | Negative and neutral words | **500ms** | Avoid-threat training caused significant decrease in negative bias | Significant differences in trait and state anxiety between groups at end of training may confound interpretation of attentional effects |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (Wells and Beevers 2009) | 4 sessions of avoid-threat or sham training over 2 weeks | Non-clinical volunteers with BDI > 8 (n=17 per group) | Visual probe | Negative and neutral faces and pictures | **Faces were presented for 3000ms, pictures for 4500ms** | Avoid-threat training caused significant decrease in negative bias | Significant differences in depressive symptoms between groups at end of training may confound interpretation of attentional effects. The stimuli were presented for an unusually long time. |
| (Amir, Beard et al. 2009) | 8 sessions of avoid-threat or control training over 4 weeks | Patients with generalised social phobia (n=22 per group) | Emotional version of Posner cueing task | Social threat and neutral words | 600ms | Avoid-threat training lead to quicker disengagement from threat words | Training and testing used different tasks and stimuli (training was achieved with a standard faces dot-probe). This study therefore demonstrates a generalisation of the training effect. Significant differences in anxious symptoms between groups at end of training may confound interpretation of attentional effects |
| (Hayes, Hirsch et al. 2010) | Single session of avoid-threat or control training | Non-clinical group (n=24 per group) with high levels of worry (PSQW > 56) | Visual probe and dichotic listening task | Threatening and neutral words, worry related and neutral texts | **750ms** (for visual probe task) | Training influenced attention (using catch trials in visual probe task) | Avoid-threat training was also associated with decreased intrusions from worrying thoughts after training. Participants were explicitly told about training contingency at the start of the study. |
| (Krebs, Hirsch et al. 2010) | Single session of avoid-threat or attend-threat training | Non-clinical group (n=32 per group) with low levels of worry (PSQW < 56) | Visual probe | Threatening and neutral words | **750ms** | Training influenced attention | Explicitly telling participants about the training contingency was found to increase the training effect. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (Najmi and Amir 2010) | Single session of avoid-threat or control training | Non-clinical participants (n=26 per group) with contamination concerns | Visual probe | Contamination relevant or neutral words | **500 ms** | Avoid-threat training reduced attention to contamination relevant words | Training also influenced the degree of behavioural avoidance of their feared objects |
| (Reese, McNally et al. 2010) | Single session of avoid-threat or control training | Non-clinical participants (n=20 per group) reporting fear of spiders. | Visual probe | Pictures of spiders or cows/birds | **500 ms** | Avoid-threat training reduced attention to spider stimuli | Training effect disappeared over following week and did not influence subjective measures |
| (Dandeneau, Baldwin et al. 2007) | Single session of positive visual search or control training | Non-clinical group (n=76/147 per study in two studies). For analysis the groups were further divided into high and low self-esteem (see note) | Emotional Stroop (study 2a) and visual probe (study 2b) | Negative, positive and neutral words (Stroop task). Happy, angry, neutral faces (visual probe task) | 1200ms for Stroop task. During visual probe task stimuli presented for **500ms** | Positive training caused decreased interference from negative words and increased interference from positive words during the Stroop task. Positive training reduced attention to angry faces in visual probe task but did not affect attention to happy faces | A significant effect of training was only found in those with low self-esteem. The influence of self esteem was assessed by splitting groups using a median split in study one and by adding self esteem as a regressor in study 2. A variety of statistical techniques were used to analyze behavioural responses in the separate studies. Stimuli were presented for 3400ms during the visual search training. |
| (Dandeneau and Baldwin 2004) | Single session of positive visual search or control training | Non-clinical group (n=49 in total). For analysis the groups were further divided into high and low self-esteem by a median split | Emotional Stroop | Negative, positive and neutral words | 1200ms | Trend level results suggest training decreased interference from negative words | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| (Baert, De Raedt et al. 2010) | 10 sessions of avoid-threat/attend-positive vs. neutral training over 10 days | Dysphoric, non-clinical group (n=23 per group) for study 1 and clinically depressed (n= 15 per group) for study 2 | Emotional Posner cueing task | Negative and positive words | 1500ms | No significant effect of training in either study | These studies employed a novel training intervention which did not appear to work |

BDI = Beck Depression Inventory, CBT = Cognitive Behavioural Therapy, GAD = Generalised Anxiety Disorder, PSWQ=Penn State Worry Questionnaire

[a] Unless otherwise stated studies used a between subjects design.

[b] Certain task designs (see figure one) allow inference as to the timing of the temporal effects of an intervention on attention. In studies which employ such a design and report significant findings the relevant duration is presented in bold.

## 1.10. Synthesis

A variety of techniques have been shown to alter emotional attention in clinical and non-clinical populations. Across all of these methodologies, interventions which improve anxiety and depression also cause attention to be preferentially deployed away from negative and towards positive information. These findings are consistent with the modification of attentional bias mediating the beneficial treatments effects of both pharmacological and cognitive interventions. The greatest amount of data is available for the pharmacological manipulation of the serotonergic system and for the use of cognitive attention training tasks. Characterisation of the attentional effects of these interventions suggests that they may target different components of the attentional control system. Specifically pharmacological modifications alter the initial deployment of attention, via an effect on the amygdala based appraisal system whereas cognitive training interventions influence attention some time later, possibly via an effect on frontal control regions.

### 1.10.1. Outstanding questions

The review presented in the previous sections identifies a number of outstanding questions concerning the mechanisms by which attentional bias may be modified. Most obviously, the neural systems altered by psychological interventions remain obscure. Where the pharmacological literature has studied the effects of interventions in a wide range of both clinical and non-clinical groups, all the neuroimaging studies of psychological interventions have been restricted to the examination of treatment effects in clinical populations.

Consequently it is not obvious whether the alterations of neural function reported in these studies result from the treatments administered or from differences in clinical status between groups (see section 1.7). Secondly, unlike the behavioural literature, the neuroimaging studies of both pharmacological and psychological interventions have tended to simply assess neural response to the presentation of emotional stimuli rather than investigating specific cognitive processes such as attention. It is thus unclear from this literature whether any of the neuroimaging effects reported can firmly be associated with alterations of attentional control per se. Lastly, if the hypothesis that psychological interventions alter attentional function primarily via an effect on frontal control systems (DeRubeis, Siegle et al. 2008; Harmer 2008) is borne out it remains unclear whether this is best conceptualised within a conflict or expectancy (see section 1.5.2) based framework.

### 1.10.2. Implications for treatment

Even in the controlled context of treatment studies neither pharmacological nor cognitive treatments of the emotional disorders produce high levels of clinical remission (for example in a large, high quality RCT DeRubeis, Hollon et al. 2005 report remission rates of 40-50%). An obvious response to these moderate remission rates is to examine the combination of psychotherapy and antidepressant medication, although studies which have done so have produced mixed results (Foa, Franklin et al. 2002; Pampallona, Bollini et al. 2004; Furukawa, Watanabe et al. 2007; Cuijpers, van Straten et al. 2009). However, all of these studies have used complex psychotherapeutic interventions rather than procedures designed to target specific cognitive systems, such as attentional bias training. An intriguing question raised by the hypothesis that antidepressant medication and attentional bias training influence distinct

attentional control systems is; how would the two treatments interact? That is, if combined, would they produce an additive effect on attention and emotion or would they interfere with each other? The behavioural and neural measures of cognition reviewed in this chapter would provide a relatively straightforward model in which this question could be addressed in a controlled, experimental manner.

Lastly, as noted above (section 1.4.3), the utility of modifying attention to emotional information may not be limited to the depressive and anxious diagnoses. There is a pressing requirement for novel treatment approaches to the symptoms of anxiety and depression in other diagnoses. This is particularly acute in the case of bipolar disorder in which symptoms of both anxiety and depression are extremely common (Kessler 1999; McIntyre and Keck 2006), but the ability to use antidepressant medication is constrained by concerns that they may induce episodes of mania (Keck, Strawn et al. 2006). Theoretically at least, attentional bias training may be a particularly useful strategy in such disorders, although direct evidence for this is lacking. A practical first step in addressing this issue would be to determine whether bipolar disorder is associated with an attentional bias to threat and whether this bias is modifiable in a controlled, experimental setting.

### 1.10.3. Conclusion

The aims of this thesis are to investigate the mechanisms by which cognitive training tasks alter attention to emotional information and to use these initial results to improve our understanding of attentional control and to guide the development of novel treatment strategies. These aims are addressed in the following chapters in which neuroimaging and

behavioural data are collected in a series of studies of both non-clinical and clinical participants. Chapter Two describes two experiments in which the extent of the attentional biases induced by different methods of training are assessed behaviourally in a non-clinical population. In Chapter Three the neural effects of attentional bias training are assessed using a task which specifically probes attention to emotional information. Chapter Four builds on the insights from this study by using a computational model to test whether expectancy based processes can predict the behavioural and neural responses to emotional stimuli. Chapter Five describes an experiment which examines the effects of combining antidepressant medication and attentional bias training on both cognitive function and emotional reactivity. Finally, Chapter Six describes a clinical study in which baseline attentional bias towards threat in a group of patients with bipolar disorder is characterised and an attempt is made to use attentional bias training to modify this bias.

# Chapter 2

# Characterising the Generalisation of Attentional Bias Training

## 2.1. Introduction

Cognitive models of anxiety and depression suggest that abnormalities in the processing of emotional information play a causal role in the aetiology of the disorders (Williams, Watts et al. 1997; Mathews and Mackintosh 1998; Mogg and Bradley 1998; Mathews and MacLeod 2005). There has been particular interest in the role of attention with a substantial body of work demonstrating an association between the emotional disorders and a tendency to deploy attention towards negative stimuli (MacLeod, Mathews et al. 1986; Williams, Mathews et al. 1996; Gotlib, Kasch et al. 2004; Gotlib, Krasnoperova et al. 2004; Fox, Russo et al. 2005; Joormann and Gotlib 2007). However, these cognitive models assert that attentional biases are more than mere correlates of anxiety and depression; rather negative attentional bias is causally implicated in the development and maintenance of the disorders (see Chapter One**)**. There has thus been increasing interest in developing experimental procedures which specifically modulate attentional processes; so called attentional bias training tasks (Mathews and MacLeod 2002; Dandeneau and Baldwin 2004; MacLeod, Koster et al. 2009; Baert, De Raedt et al. 2010). While research using such tasks has demonstrated their utility both as an experimental tool (MacLeod, Rutherford et al. 2002) and a clinically useful intervention in their own right (Amir,

Beard et al. 2009; Schmidt, Richey et al. 2009; Wells and Beevers 2009), a number of outstanding questions on the nature of the induced biases remain. The first question concerns the degree to which induced attentional biases generalise. For example, although MacLeod and colleagues (2002) confirmed that a training task altered attentional bias, assessment of attention was limited to a word based visual probe task. Since the training procedure employed in this study was itself a word based visual probe task this leaves open the possibility that the attentional effect of training may be limited only to the same stimulus type as that used in the induction task. Clearly, if it is to be argued that training alters the generalised attentional bias found in the emotional disorders (Mathews and MacLeod 1985; Bradley, Mogg et al. 1997; Bradley, Mogg et al. 1997; Yiend and Mathews 2001; Gotlib, Kasch et al. 2004), rather than simply altering the response to a limited set of stimuli, it is necessary to demonstrate that any effect of training is not confined only to the exact training material. That is, demonstration of generalisation to a novel (in this case, non-word) emotional stimulus type is required. A second, directly practical question, concerns the relative efficacy of the various training regimes; put simply, some attentional bias training paradigms fail to produce reliable alterations of attentional function. For example, in two large studies Baert and colleagues describe a novel training paradigm based on Posner's spatial cueing task (Posner 1980) which failed to alter attentional function in either a non-clinical group or a group of depressed patients (Baert, De Raedt et al. 2010). The obvious conclusion from these results is that it cannot be assumed that every training intervention will be effective.

The essential requirement of the current chapter was therefore to ensure that an effective version of the training task was developed which demonstrably altered the deployment of attention to emotional information. In order to achieve this two studies were performed which assessed the behavioural effects of a word based and a face based training paradigm in non-clinical populations. The studies also sought to characterise the nature of any induced biases by assessing the extent of bias generalisation in two dimensions; firstly, as described above, the studies tested whether a bias induced using one stimuli type (such as words) would alter attention to another type of stimuli, secondly, the generalisation of the training effect across time was assessed by measuring the deployment of attention both at the same duration as training (500 ms) and at a shorter duration (100 ms) which has more commonly been associated with the effect of pharmacological manipulations (see Chapter One).

The key prediction of the current studies was that attentional bias training would induce a bias of attention to emotional stimuli when assessed using the same stimuli type, that is the central finding of MacLeod at al. (2002) would be replicated. A second prediction was that the effects of attentional bias training would generalise to influence measures which used different types of stimuli and those in which the stimuli were presented for shorter durations.

## 2.2. General Methodology

### 2.2.1. General Overview

Two randomised experimental studies were performed which compared the behavioural effects of "attend-negative" with "avoid-negative" attentional bias

training (see Figure 2.1 and sections 2.3.1.2 and 2.4.1.2 for descriptions of the training tasks). Attend-negative training was designed to encourage participants to preferentially deploy attention towards threatening stimuli where avoid-negative training was designed to have the opposite effect. The studies were identical in structure except that in study one a word based training procedure was used whereas study two employed a face based task. In both studies, following randomisation into either the attend-negative or avoid-negative conditions, participants completed the attentional bias training regime. Participants then undertook a test of attentional deployment using a standard visual probe task which displayed stimuli of the same type as the training task. (i.e. words were used after word training and faces after face training). Participants then completed a "booster" session of training and finally a second test visual probe task, this time using the opposite type of stimuli (i.e. faces after word training and vice versa). In both studies the training stimuli were presented for 500 ms whereas attentional bias was assessed at both 500 and 100 ms.

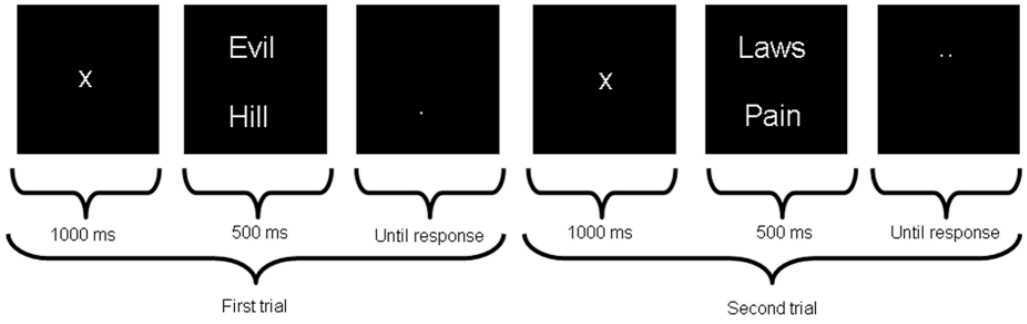



**Figure 2.1: Two example trials from the "avoid-negative" version of the word based training task. The task is based on the visual probe paradigm, described in chapter one, except that the position of the probe is controlled. In each trial a negative and neutral word are presented. A probe, to which the participant must respond, follows the words. In the avoid-negative condition (shown) the probe always replaces the neutral word. The attend-negative condition is identical in every respect except that the probe is found in the location of the negative word. Over the course of many trials participants are thought to develop the habit of attending to the stimulus type which predicts the probe location.**

### 2.2.2. Participants

Participants were 24 healthy individuals per study (48 in total) recruited from the local community. Inclusion criteria were age 18-60 and speaking English as a first language; exclusion criteria included current neurological illness and the use of any medication other than the hormonal contraceptive pill. In addition, all participants were screened to exclude current or past axis I psychiatric illness or alcohol/substance misuse using the Structured Clinical Interview for the DSM-IV (SCID; Spitzer et al. 2002). The studies were reviewed and approved by the Central University Research Ethics Committee of the University of Oxford. Written informed consent was obtained from all participants.

## 2.3. Study 1

### 2.3.1. Methods

#### *2.3.1.1. Self report mood and anxiety measures*

Baseline measures of depressive symptoms were obtained using the Beck Depression Inventory (BDI; Beck, Ward et al. 1961), an extensively used 21 item self report measure which categorises the severity of depressive symptoms over the previous week. The severity of each symptom (e.g. feelings of guilt) is rated 0-3, with the summary statistic being the sum of all individual items, resulting in a minimum total score of 0 and a maximum, indicating high levels of depressive symptoms, of 63. Anxious symptoms were assessed using the trait subscale of the Spielberger State-Trait Anxiety Inventory (trait-STAI; Spielberger, Gorsuch et al. 1983). This self report measure asks participants to indicate the extent to which 20 anxiety related

statements (such as “I have disturbing thoughts”) describes the way they feel generally. Participants answer using a four point scale (anchored at “not at all” and “very much”) resulting in a minimum score for each item of 1 and maximum of 4. The scale is scored (after accounting for reversed items) by summing across the individual items resulting in a minimum score of 20 and a maximum score, reflecting high levels of anxiety, of 80. The direct effect of training on momentary mood was assessed using visual analogue scales (VAS) measuring “happy” and “sad” (each anchored at “not at all” and “very”) as well as the state version of the STAI (state-STAI) which is of an identical structure to the trait-STAI with the exception that it assesses how participants feel at that point in time rather than generally. All momentary mood measures were completed immediately before and after training.

*2.3.1.2. Word based attentional bias training task*

The training task (see Figure 2.1) was designed to replicate that published by MacLeod et al. (2002) and used identical materials. It is based on the standard visual probe task in which two words are presented concurrently. One of the words is negative (e.g. “coffin”) and the other neutral (e.g. “edited”). Both words then disappear and are replaced by a probe (one or two dots) in the location of one of the words. In the attend-negative group the probe is always in the position of the negative word whereas in the avoid-negative group the probe is always in the position of the neutral word. The rationale behind the training task is that by repeating the procedure over more than 500 trials a habit is formed such that participants begin to automatically and consistently attend to the stimulus type where the probe is found (Hertel and Brozovich 2010). Thus the attend-negative group, in which the probe replaces the negative word, are believed to develop a tendency to attend to negative

stimuli generally (a negative attentional bias) whereas the avoid-negative group, in which the probe replaces the neutral word, develop a tendency to direct their attention away from negative stimuli generally (a relatively positive attentional bias).

The task was presented on a computer running Eprime software (Version 1.2.1.68; Psychological Software Tools, Pittsburgh, PA). Each trial began with a central fixation cross for 1000 ms followed immediately by a word pair. One word was presented above and one below the previously presented fixation cross (each word subtended a vertical visual angle of ≈1.6º, the words were separated by ≈2.5º). The word pairs used were identical to those used by Macleod et al. (see appendix of MacLeod, Rutherford et al. 2002) comprising two sets of 48 word pairs; in each pair one word is negative, the other neutral. Training for any particular individual used words from only one set, counterbalanced between participants (the other set, which the individual had not seen was then used in the test visual probe task later on). Words were presented for 500 ms and were then immediately replaced by a probe (either "." or ".."). The negative words were presented in the upper and lower locations with equal frequency. Participants responded by button press to indicate the probe type. The task consisted of 576 trials presented in pseudorandom order with three rest periods which occurred every 150 trials.

*2.3.1.3. Word visual probe; assessment of the induced attentional bias.*

In the word visual probe task, participants were presented with a set of word pairs to which they had had no prior exposure. The spatial arrangement of the word stimuli was identical to the training task although the probes were now equally likely to replace either the negative or the neutral word thereby allowing assessment of the

induced bias. Word pairs were presented for either 100 ms or 500 ms in two blocks of 60 trials each (order of blocks was counterbalanced between participants). That is, each participant viewed one block in which the words were presented for 500 ms and one block in which they were presented for 100 ms. The visual probe has been described as providing a "snap-shot" of attention at the instant the words disappear (Cooper and Langton 2006), thus presenting the words for two different durations was intended to allow analysis of attentional deployment at these two time points.

*2.3.1.4. Booster training*

The booster session consisted of a subset of 144 trials of the training task. The rationale for the booster was that in the preceding visual probe task the contingency between emotional word and probe location was necessarily removed, possibly eroding any induced bias. It was hoped to ameliorate this by using a brief refresher training session prior to the next assessment task.

*2.3.1.5. Face visual probe; assessment of generalisation of the induced attentional bias.*

Stimuli were 16 different characters (8 female) taken from two previously validated sources (Matsumoto and Ekman 1988; Lundqvist, Flykt et al. 1998). Each character displayed either a fearful or neutral expression. Faces were presented in pairs with each pair consisting of one neutral and one fearful face. The pictures were displayed subtending a vertical visual angle of ≈3.7º and were separated by ≈0.7º. The timings of stimuli presentation were identical to the word based visual probe task in that the stimuli were presented in two blocks with faces being displayed for 100 ms or 500 ms within a block. However, as it has been suggested that pictures of facial affect can provide a more salient and intense emotional stimulus than words (Mogg and Bradley

1998), it was conceivable that any induced bias may be overwhelmed by existing attentional biases which have previously been demonstrated even in non-anxious participants using face stimuli (Mogg and Bradley 1998; Cooper and Langton 2006). Morphing software was therefore used to create a range of intensities of fearful faces with the aim of improving the sensitivity of the task to detect an induced bias; a strategy supported by the finding that attentional response to emotional faces depends on the intensity of the expression displayed (Wilson and MacLeod 2003). Stimuli were created at 25% increments along the continuum from 100% fearful to 100% neutral (i.e. 100%, 75%, 50%, 25%, 0% fearful). Each intensity was presented 20 times giving a total of 100 trials per block; and 200 trials in the overall task.

*2.3.1.6. Statistical analysis*

As in MacLeod et al. (2002) median reaction times (RT) from correct trials were used as the measure of central tendency because of its resistance to the effect of outliers. Thus median RTs were entered into a mixed model analyses of variance (ANOVA) with the between-subjects factor of group (attend-negative vs. avoid-negative), and within-subjects factors of probe validity (probe replaces negative word vs. probe replaces neutral word) and duration (100 ms vs. 500 ms). For analysis of the faces visual probe task, intensity of fearful face (4 levels 25-100%) was added as a within-subjects factor. In order to graphically demonstrate results a "vigilance score" was calculated (Wilson and MacLeod 2003) by subtracting RT when the probe replaces the negative stimuli from RT when the probe replaces the neutral stimuli [i.e. $RT_{(neut)} - RT_{(neg)}$]. Derived thus, a higher positive vigilance score reflects a bias towards the negative stimuli, and a more negative score reflects bias away from the negative stimuli. Differences between groups for the baseline questionnaire and demographic

measures were assessed using independent t-tests, with differences in gender being assessed using a Chi-Square test. Changes in momentary mood were analysed using split model ANOVAs in which time (before training, after training) was a within subject factor and group (avoid-negative, attend-negative training) a between subject factor. For the VAS analysis a difference score ("sad"-"happy") was used as an index of depressive symptoms.

## 2.3.2. Results

### *2.3.2.1. Demographics*

As can be seen from Table 2.1, there was no significant differences between experimental groups in terms of age, gender, trait anxiety (trait-STAI), BDI or initial anxiety (state-STAI) or depressive symptoms (VAS).

**Table 2.1: Demographic details and baseline measures for the participants in both studies.**

| | **Study 1** | | | **Study 2** | | |
|---|---|---|---|---|---|---|
| | Avoid-negative training (n=12) | Attend-negative training (n=12) | *p* | Avoid-negative training (n=12) | Attend-negative training (n=12) | *p* |
| **Age** mean (sd) | 21.4 (2.9) | 24.3 (6.3) | 0.2 | 25.9 (12.3) | 21.4 (2.8) | 0.2 |
| **Female:Male**[a] | 8:4 | 7:5 | 0.7 | 6:6 | 6:6 | 1.0 |
| **Trait anxiety** mean (sd) | 31.2 (7.1) | 31.9 (5.1) | 0.8 | 33 (6.7) | 32.6 (4.3) | 0.9 |
| **BDI** mean (sd) | 2.8 (3.1) | 2.6 (1.4) | 0.9 | 3.1 (3.9) | 2.7 (1.9) | 0.7 |
| **State Anxiety** mean (sd) | 27.1 (5.2) | 28.8 (4.9) | 0.4 | 31.4 (4.9) | 28.4 (5.5) | 0.2 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **VAS (sad-happy)** mean (sd) | -6.3 (1.3) | -5.3 (1.8) | 0.2 | -5.6 (1.6) | -6.5 (2.1) | 0.3 |

[a] Comparison between groups conducted using Chi-Square test, all other comparisons used independent t-tests

*2.3.2.2. Did the word training task induce an attentional bias to novel words?*

A mixed model ANOVA of the word visual probe data (see Figure 2.2a and Table 2.3a) revealed a significant three way interaction of group x duration x validity [$F(1,22) = 6.1$, $p = 0.02$]. As illustrated in Figure 2.2a this interaction was driven by a significant effect for the 500 ms exposure [$F(1,22) = 4.2$, $p = 0.05$], with no effect at 100 ms [$F(1,22) < 1$, $p = 0.3$]. Specifically, the effect at 500 ms was in the direction predicted by group, such that the negative group showed a greater attentional bias towards the negative words than did the neutral group. This result replicates that reported by MacLeod and colleagues (2002) in that the bias induction procedure was effective at the longer exposures and that it altered attention to novel words.

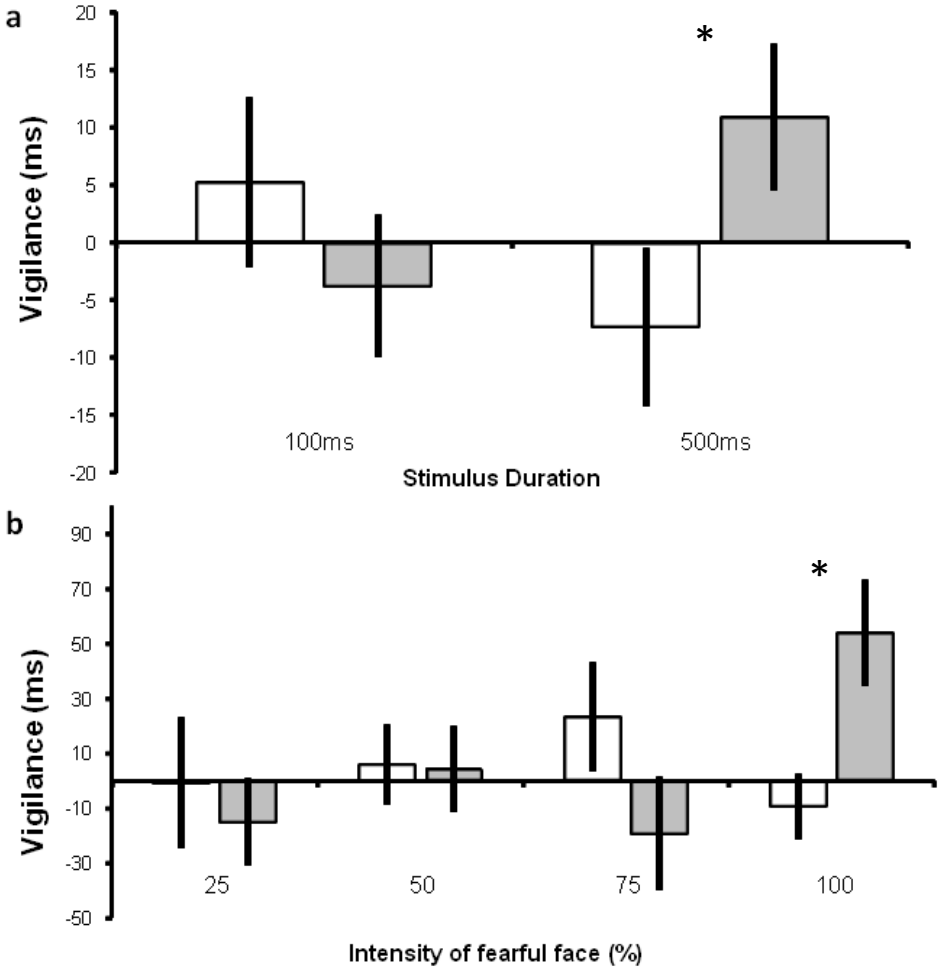



**Figure 2.2: The attentional effects of word based training. (a) Vigilance towards novel words. Vigilance is calculated such that a positive number indicates an attentional bias towards the negative stimulus, and a negative number indicating an attentional bias away from the negative stimulus. Data from the 100 and 500 ms blocks are arranged on the x axis. (b) Vigilance towards emotional faces during the 500 ms block. The intensity of the morphed negative face is arranged along the x axis. White bars = avoid-negative group, grey bars = attend-negative group. Error bars = SEM, *=p≤0.05.**

*2.3.2.3. Did the attentional bias, induced using words, generalise to faces?*

A mixed model ANOVA of the face visual probe (Table 2.3a) revealed a significant group x duration x intensity x validity interaction [$F(3,66) = 3.3$, $p = 0.03$]. Decomposition of this finding revealed a significant group x intensity x validity interaction at 500 ms [$F(3,66) = 3.2$, $p = 0.03$]. The interaction at 100 ms did not

reach significance [$F(3,66) = 2.3$, $p = 0.09$]. Figure 2.2b displays the findings at 500 ms and indicates that the effect was driven by a significant difference between groups at the most intense level of fearful expression [$F(1,22) = 8.6$, $p = 0.03$ (Bonferroni corrected for multiple comparisons)]. The effect of training at other intensities was not significant [all corrected *p's* > 0.5] and there were no significant differences at any intensity when the faces were presented for 100 ms [all corrected *p's* > 0.1]. Again, the effect of attentional bias training found at 500 ms was in the expected direction with the negative group displaying an increased attentional bias for fearful faces. This finding suggests that the effects of word based training generalised to a faces measure, although the effect of training was only evident when prototypical pictures of fear were presented for 500 ms.

**Table 2.2: Changes in anxious (state-STAI) and depressive (sad-happy VAS) symptoms across training in study one. All figures are presented as mean (standard deviation)**

| | **Avoid-negative training** | **Attend-negative training** |
|---|---|---|
| **State-STAI pre training** | 27.2 (5.2) | 28.8 (4.9) |
| **State-STAI post training** | 33.8 (6.6) | 32.1 (6.4) |
| **Sad-Happy VAS pre training** | -6.3 (1.3) | -5.3 (1.8) |
| **Sad-Happy VAS post training** | -5.8 (2.4) | -4.6 (2.3) |

STAI= Spielberger State-Trait Anxiety Inventory. VAS= Visual Analogue Scale

*2.3.2.4. Did word training differentially influence state anxiety or mood?*

Although, by design, the training task is intended to directly alter attentional function, an alternative interpretation of these effects is that training acts as a form of mood induction (e.g. Velten 1968) with the altered mood state leading to the observed

attentional effects. This interpretation was tested by comparing the effects of training on changes in the mood VAS and the state-STAI across training (see Table 2.2). In both groups state anxiety increased across training [$F(1,22) = 20.6$, $p < 0.001$; increase in avoid-negative group = 6.6; attend-negative group = 3.3]. However, this effect was not significantly influenced by group membership [$F(1,22) = 2.4$, $p = 0.13$] and, contrary to the mood induction hypothesis, was numerically greater in the avoid-negative group. Although there was a tendency for depressive symptoms, as measured by the sad-happy VAS measures, to increase across training (change in avoid-threat training = 0.5; attend-threat training = 0.7), this increase was not significant [$F(1,22) = 2.4$, $p = 0.12$] and there was no time x group interaction [$F(1,22) < 1$, $p = 0.8$]. These results are consistent with the findings reported by MacLeod and colleagues (2002) and suggest that the influence of training on attention cannot be accounted for by a simple mood induction effect.

**Table 2.3: Raw group reaction times from Study One (a) and Study Two (b). Mean (standard deviations) are reported for each trial type (e.g. the reaction time for neutral words gives the results for those trials in which the probe replaced the neutral word) and each group to the nearest millisecond**

a.

| | **Avoid-negative training** | | | | | | | | **Attend-negative training** | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Duration of Stimuli** | **100** | | | | **500** | | | | **100** | | | | **500** | | | |
| **Neutral Words** | 458 (51) | | | | 431 (49) | | | | 468 (41) | | | | 464 (55) | | | |
| **Negative Words** | 453 (48) | | | | 439 (60) | | | | 472 (44) | | | | 453 (54) | | | |
| **Intensity of Negative Face (% fear)** | **25** | **50** | **75** | **100** | **25** | **50** | **75** | **100** | **25** | **50** | **75** | **100** | **25** | **50** | **75** | **100** |
| **Neutral Faces** | 505 (77) | 519 (91) | 503 (70) | 512 (57) | 512 (118) | 527 (114) | 519 (96) | 526 (107) | 555 (81) | 558 (73) | 562 (85) | 566 (87) | 543 (95) | 553 (81) | 559 (79) | 606 (102) |
| **Negative Faces** | 509 (82) | 502 (78) | 519 (96) | 510 (79) | 513 (120) | 521 (116) | 496 (108) | 535 (105) | 571 (94) | 591 (92) | 559 (79) | 570 (73) | 558 (71) | 548 (86) | 578 (112) | 552 (73) |

b.

| | Avoid-negative training | | | | | | | | Attend-negative training | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Duration of Stimuli** | **100** | | | | **500** | | | | **100** | | | | **500** | | | |
| **Neutral Words** | 517 (86) | | | | 504 (124) | | | | 539 (79) | | | | 503 (63) | | | |
| **Negative Words** | 512 (79) | | | | 492 (123) | | | | 542 (91) | | | | 503 (64) | | | |
| **Intensity of Negative Face (% fear)** | **25** | **50** | **75** | **100** | **25** | **50** | **75** | **100** | **25** | **50** | **75** | **100** | **25** | **50** | **75** | **100** |
| **Neutral Faces** | 496 (60) | 493 (63) | 490 (68) | 484 (63) | 451 (70) | 476 (81) | 480 (70) | 474 (80) | 494 (54) | 468 (51) | 484 (52) | 483 (47) | 476 (110) | 465 (55) | 459 (46) | 467 (67) |
| **Negative Faces** | 495 (53) | 487 (71) | 494 (78) | 496 (57) | 473 (74) | 461 (70) | 482 (74) | 474 (85) | 477 (71) | 480 (51) | 488 (68) | 497 (70) | 448 (57) | 466 (58) | 454 (47) | 465 (80) |

### 2.3.3. Summary

The key results reported above are that a word based training task induced the expected bias of attention to novel words in a group of non-clinical participants and that this bias generalised to the extent that it remained apparent when attention to prototypical fearful faces was assessed. In contrast, the induced bias appears only to manifest at the longer duration (500 ms) with no significant effect of training being found at the shorter (100 ms) duration. The bias is unlikely to be accounted for by a mood induction effect of training as no differences between the groups were found on measures of anxiety or depressive symptoms. In study two a similar assessment is made of a face based training task.

## 2.4. Study 2

### 2.4.1. Methods

#### *2.4.1.1. Overview*

The design and analysis of this study replicated that described above in almost all aspects. The essential change made in the current study is that a face based training paradigm was used in place of the word based training assessed in study one. Additional differences in the task parameters are summarised below.

#### *2.4.1.2. Face based attentional bias training task*

The structure, duration and procedure of this task were identical to the word based training. The stimuli used in the task were identical to those used in the previous faces visual probe assessment task; although stimuli were split into two groups (of 8 characters each) with one group of stimuli being used in training and the other in testing. Only the prototypical faces

(i.e. 100% fearful) were used in training although the full range of morphed faces was used in testing.

*2.4.1.3. Assessment tasks*

The assessment visual probes were identical to that described for study one with the exception that their order was reversed (the faces visual probe was performed first) and that pictures of only 8 characters were used in the faces task as the remaining 8 characters had been used in training.

## 2.4.2. Results

*2.4.2.1. Demographics*

As can again be seen from Table 2.1, there was no significant differences between experimental groups in terms of age, gender, trait anxiety, BDI or initial anxiety or depressive symptoms.

*2.4.2.2. Did the face training task induce an attentional bias to novel faces?*

Omnibus ANOVA of the faces data (see Figure 2.3 and Table 2.3b) revealed a significant interaction of group x intensity x validity [$F$ (3,66) = 3.6, $p$ = 0.04], which was not further qualified by an interaction with duration [$F$ (3,66) < 1, $p$ = 0.6]. Post hoc tests indicated a trend effect of group, only at the lowest intensity of fearful face (25%) [$F$ (1,22) = 5.8, $p$ = 0.08 (corrected for multiple comparisons)] with the other differences being non-significant. Again this effect is in the predicted direction with the negative group displaying a more marked attentional bias to fearful faces.

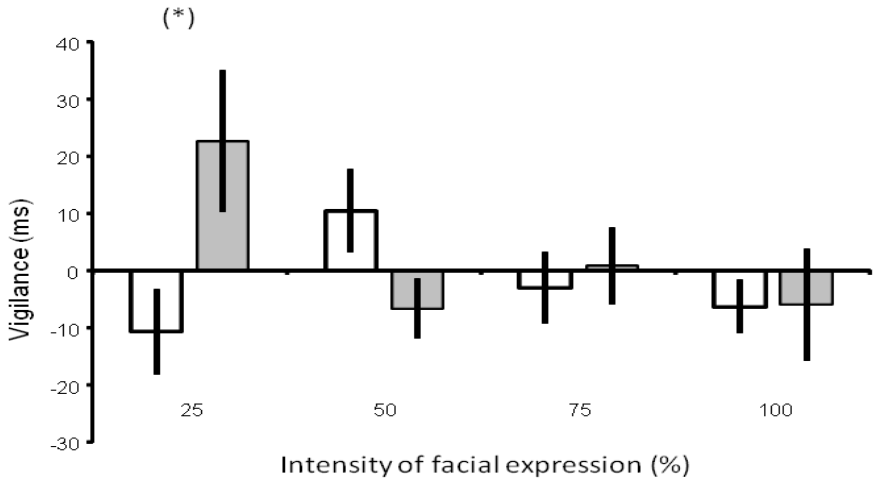



**Figure 2.3: Attentional effects of face training, vigilance to novel faces (regardless of duration). White bars = avoid-negative group, grey bars = attend-negative group. Error bars = SEM. (*) = trend level difference between groups.**

*2.4.2.3. Did the attentional bias, induced using faces, generalise to words?*

There was no effect of group on attention to words (all *F's* < 1) and thus no evidence that the attentional bias induced by faces generalised to words (Table 2.3b).

*2.4.2.4. Did face based training differentially influence state anxiety or mood?*

There was a tendency for state anxiety (see Table 2.4) to increase across face training [$F$ (1,22) = 3.8, $p$ = 0.06; increase in avoid-negative group = 1; attend-negative group = 2.8], however this was not modified by group [$F$ (1,22) < 1, $p$ = 0.4]. There was no significant change in depressive symptoms across training [$F$ (1,22) = 1.4, $p$ = 0.25; change in avoid-negative training = -0.9; attend-negative training = -1.2] and no group x time interaction [$F$ (1,22) < 1, $p$ = 0.66] confirming that no mood induction-like effects occurred during training.

**Table 2.4: Changes in anxious (state-STAI) and depressive (sad-happy VAS) symptoms across training in study two. All figures are presented as mean (standard deviation)**

| | Avoid-negative training | Attend-negative training |
|---|---|---|
| **State-STAI pre training** | 31.4 (4.9) | 28.4 (5.4) |
| **State-STAI post training** | 32.4 (5.3) | 31.2 (6.3) |
| **Sad-Happy VAS pre training** | -5.6 (1.6) | -5.1 (3) |
| **Sad-Happy VAS post training** | -6.5 (2.1) | -6.3 (2) |

STAI=Spielberger State-Trait Anxiety Inventory. VAS=Visual Analogue Scale

### 2.4.3. Summary

The effect of face based training on emotional attention was marginal. An effect of training on attention to novel faces was found which was modified by the intensity of the negative face (but not by the duration for which the faces were presented). This appears to have resulted from an effect of training which was evident at the lowest stimulus intensity, although when correcting for multiple comparisons this effect survived only at a trend level. Further there was no evidence that the effect of face training generalised to words.

## 2.5. Discussion

The studies in the current chapter successfully replicate the pivotal finding of MacLeod at al. (2002), demonstrating that word based attend-negative versus avoid-negative training contingencies differentially modulate attentional bias to novel word stimuli. Further, by analysing the generalisation of training the current studies provide a more fine grained assessment of the induced attentional effects than previously reported. Critically, the effects of word based training were not limited just to other words but were also apparent using a test of attention to face stimuli.

### 2.5.1. The basic attentional bias training effect

The central task of the current chapter was to ensure that the basic attentional bias training effect could be replicated in the current laboratory. The results from study one provide assurance on this as the word based attentional bias training task produced clear effects on the deployment of attention when assayed using two different measures of attention (effect size when comparing mean vigilance scores of the attend-threat and avoid-threat groups for words presented for 500 ms; $d = 1.3$). In contrast, the effects of the face based training were less impressive; there was no evidence for the generalisation of the training effect to words and, while there was some evidence that it did alter attention to novel faces, even here the effect was not as clear cut as that found following word based training as no post hoc test survived correction for multiple comparisons. While a number of previous studies have demonstrated that attentional function may be modified using face based tasks (Dandeneau, Baldwin et al. 2007; Li, Tan et al. 2008; Amir, Beard et al. 2009), others have either not reported any attentional effects (Schmidt, Richey et al. 2009) or have found limited (Eldara, Ricona et al. 2008) or no effect (Baert, De Raedt et al. 2010). In contrast, all previously reported word based training studies (MacLeod, Rutherford et al. 2002; Hazen, Vasey et al. 2008; Wadlinger and Isaacowitz 2008; Amir, Beard et al. 2009; See, Macleod et al. 2009) report significant effects of training. Thus, while it would seem to be possible to use faces to induce attentional biases the inconsistency in the training effect appears to be greater than when word stimuli are used.

### 2.5.2. The generalisation of attentional bias training effects

The negative attentional biases reported in clinical anxiety and depression deploy across word, face and picture stimuli (Mathews and MacLeod 1985; Bradley, Mogg et al. 1997;

Yiend and Mathews 2001; Gotlib, Kasch et al. 2004). Similarly, the bias induced by word based training in study one affected more than one stimulus type, and critically a type not involved in the training procedure. As argued in the introduction, demonstration of such generalisation is crucial to support the proposal that training alters attention to the emotional content of stimuli rather than producing a limited effect simply on a single class of stimulus. Interestingly, the effects of word based training did not seem to generalise across stimulus duration; while the deployment of attention was assessed at both relatively early (100 ms) and late (500 ms) durations (Cooper and Langton 2006) an effect of word based training was only found at the longer duration. As reviewed in Chapter One, this finding is consistent with the other published effects of training which, when compared to pharmacological interventions, appear to alter the later deployment of attention. However, interpretation of this result is constrained by the fact that the training tasks used in the current studies presented stimuli for 500 ms -- thus the later onset of the attentional bias may simply reflect this timing parameter of the training task. It is possible that if the training task had been performed using a 100 ms presentation of stimuli an early effect would manifest at testing.

### 2.5.3. Limitations

In Chapter One it was suggested that attentional bias training may provide a model of at least one of the mechanisms involved in more complex psychological interventions such as CBT. However, the current study compared avoid-negative training and attend-negative training, either or both of which may actively influence attentional function. While this design provides the most sensitive measure of the behavioural consequences of attentional bias training, it cannot discriminate whether the observed affects result from the attend-negative training, the avoid-negative training or both together. Separation of these possibilities would

require an extra, control group to be administered a neutral form of training. As the avoid-negative training is predicted to be therapeutic in anxiety and depression (Amir, Beard et al. 2009; MacLeod, Koster et al. 2009; Schmidt, Richey et al. 2009; See, Macleod et al. 2009; Wells and Beevers 2009) an interesting next step would be to assess the effects of this form of training in comparison to such a neutral control condition.

There are several further limitations to these studies. Firstly, attentional bias to the spatial locations of only word and face stimuli was assessed. However, there is also evidence linking the emotional disorders to negative attentional biases towards the location of non-visual stimuli (e.g. using the dichotic listening task; Mathews and MacLeod 1986; Hayes, Hirsch et al. 2010) and when using non-spatial measures of attention (e.g the attentional blink; Fox, Russo et al. 2005) as well as when using other types of visual stimuli (e.g emotional scenes; Yiend and Mathews 2001). It would therefore be interesting to explore the generalisation of induced biases in all of these other domains. Secondly, the current studies used only a single experimental session of training. It is possible that the attentional effects of repeated training may be more pronounced than that found following a single sessions; in particular the equivocal effects of face training found in the current study may have been more pronounced had training been repeated. However, the finding that word based training did produce a clear alteration of attentional function within session indicates that single session studies do provide a useful tool for probing the acute effects of training. A related concern is that the negative results from the face-based training study may be a consequence of the study not being sufficiently powered to detect the relevant effects. Thus, while an identical sample size was sufficient to detect the impact of word training, future studies, with

larger sample sizes, may be able to provide a more complete assessment of the effects of face-based training.

### 2.5.4. Conclusion

In conclusion, the current chapter has demonstrated that a single session of word based training can modify attentional bias to two types of emotional stimuli in a non-clinical population. These results form the basis for the study reported in the following chapter which investigates the neural mechanisms underlying the training effect. Specifically, an identical word based training procedure is deployed in a non-clinical population and the effects of training on neural activity are then assessed using a task in which pictures of facial affect are presented.

# Chapter 3

# Can Alteration of Prefrontal Function Account for the Effects of Attentional Bias training? An fMRI Study

## 3.1. Introduction

As reviewed in Chapter One, a range of evidence indicates that negative attentional biases are causally linked to anxious and depressive symptoms and that these biases may be modified using either pharmacological or psychological strategies.

Neural models (Desimone and Duncan 1995; Pessoa and Ungerleider 2005; Vuilleumier 2005; Bishop 2007) of attentional control suggest that two biasing signals influence the deployment of attention to emotional stimuli. An amygdala based system produces a signal which promotes the deployment of attention towards salient stimuli (Whalen, Rauch et al. 1998; Davis and Whalen 2001; Pessoa, Padmala et al. 2005; Phelps and Le Doux 2005; Lim, Padmala et al. 2009), whereas a more flexible response is associated with a second signal, originating in areas of the prefrontal cortex (including the rostral anterior cingulate cortex [rACC] and the lateral prefrontal cortex [lPFC]; MacDonald, Cohen et al. 2000; Bishop, Duncan et al. 2004a; Etkin, Egner et al. 2006). The prefrontal control signal is believed to play a particularly important role when conflicting demands are made on attention. Both

kinds of biasing signal are thought to harness processing resources in favour of their preferred, and at the expense of their less preferred, stimuli. In neural terms, increased attention to a stimulus, generated by either the amygdala or prefrontal cortical systems, is associated with increased activation of the relevant sensory and association cortices in response to that stimulus (Vuilleumier and Driver 2007). Interventions which modify emotional attention may thus plausibly be mediated by alteration of the function of either the amygdala or the prefrontal biasing signals; the effects of the interventions on attention would be predicted to be reflected in altered sensory and association cortex activation to emotional stimuli.

Direct experimental evidence indicates that antidepressant medications reduce amygdala activation to threatening stimuli (Sheline, Barch et al. 2001; Fu, Williams et al. 2004; Del-Ben, Deakin et al. 2005; Harmer, Mackay et al. 2006; Murphy, Norbury et al. 2009) and increase visual association cortex response to positive stimuli (Norbury, Mackay et al. 2007), suggesting that these drugs may alter attentional habit via an effect on stimulus appraisal rather than on higher order control processes. It has been suggested that psychological treatments, in contrast, are likely to work through changes in the frontal control systems (DeRubeis, Siegle et al. 2008; Harmer 2008). While this seems plausible, the complexity and variability of formal evidence-based psychotherapies such as cognitive behavioural therapy (CBT) complicate the interpretation of their effects in controlled experimental trials. It appears more logical to study experimentally the mechanisms of their component procedures. Using this approach explicit methods of emotional reappraisal have been demonstrated to be associated with alteration in prefrontal function (Ochsner, Ray et al. 2004). However, there is little evidence regarding the mechanisms by which habitual attentional bias may be

influenced. Accordingly, the current chapter investigates the mechanisms by which cognitive training (Mathews and MacLeod 2002) alters attentional control functions as reflected by BOLD fMRI signal. An identical training procedure was used to the word based training described in Chapter Two which demonstrably altered a behavioural measure of attentional function. The primary hypothesis was that the neural mechanism by which training leads to changes in attentional bias is alteration of rACC and lPFC function (DeRubeis, Siegle et al. 2008; Harmer 2008). A related prediction was that these changes in frontal function would be associated with secondary changes in visual sensory association cortex which provide a neural measure of attention (Vuilleumier 2005).

## 3.2. Methods and Materials

### 3.2.1. Participants

A total of 29, right-handed, native English speaking, healthy participants recruited by advertisement from the general population were randomly assigned to either "attend-negative" or "avoid-negative" training conditions (see section 3.2.3). Participants were screened to exclude current or previous axis I psychiatric disorder or alcohol/substance misuse using the Structured Clinical Interview for the DSM-IV (Spitzer, Williams et al. 2002). Participants were also excluded if they were taking any psychoactive medication, had any significant neurological condition, were familiar with any of the tasks or stimuli used in the study or had any contraindication to undertaking an MRI scan. All participants provided written informed consent to the study which had been approved by an Oxfordshire Research Ethics Committee. Immediately following the attentional bias training task (attend-negative= 14, avoid-negative =15) the effects of the training procedure was assessed using an fMRI paradigm.

### 3.2.2. Questionnaire Measures

Participants completed questionnaire assessments of depressive (Beck Depression Inventory [BDI]; Beck, Ward et al. 1961) and anxious symptoms (Trait subscale of the State-Trait Anxiety Inventory (trait-STAI); Spielberger, Gorsuch et al. 1983) before scanning. State anxiety and mood were also assessed before and after completion of the training task (using both state-STAI and visual analogue scales (VAS) measuring "happy" and "sad") to monitor whether the training task induced any changes in mood.

### 3.2.3. Attentional Bias training Task

The attentional bias training task was identical to the word based training described in Chapter Two, study one. Briefly, this involved 576 trials in which a negative and neutral word were presented for 500 ms. The word pair was then followed by a probe to which the participant had to respond. For the avoid-negative group the probe was always in the location of the neutral word, whereas for the attend-negative group the probe was always in the position of the negative word (Figure 2.1). Before training all participants completed a practice version of the imaging task (see section 3.2.4).

### 3.2.4. Imaging Task

Immediately following the training session, the effects of training on neural activity was assessed with a task (see Figure 3.1) which was adapted from Pessoa and colleagues (Pessoa, Padmala et al. 2005). This task was selected as it allows the influence of both the emotional content of stimuli and the direction of attention to be discriminated. Further, as the task presents only one emotional image at a time it was felt likely that behavioural performance on the task would not be affected by the attentional bias training undertaken; using a behaviourally insensitive task allows a more straightforward assessment of the imaging

results as it removes possible behavioural confounds. Following a centrally presented fixation cross a picture of a face (Ekman and Friesen 1976; Matsumoto and Ekman 1988; Lundqvist, Flykt et al. 1998; Tottenham, Tanaka et al. 2009) flanked by two bars (vertical visual angle of face picture ≈ 4 °, eccentricity of bars ≈ 6 °) was presented for 200 ms. Manipulation of the affective quality of the stimuli was achieved by presenting either fearful or neutral faces (NB only 100% fearful faces were used in scanning as it was only at this intensity that a behavioural effect of training was found in Chapter Two). The direction of attention of participants was manipulated using sequential blocks of twenty trials during which participants were instructed to respond by button press to either the gender of the face (i.e. requiring that attention is focused on the face) or to whether the flanking bars were aligned (i.e. requiring that attention is directed away from the face).

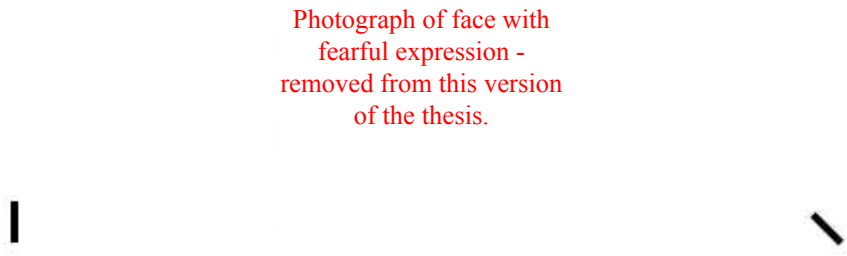



**Figure 3.1 Behavioural task completed during the scan, example trial (Pessoa, Padmala et al. 2005). On each trial a central face was presented flanked by two bars for 200 ms. The face displayed either a fearful (shown) or neutral expression. Participants were instructed to report either the gender of the face (i.e. requiring attention to be focussed on the face) or whether the flanking bars were aligned (i.e. requiring attention to be directed away from the face).**

The overall structure of the task was thus factorial with two levels of emotion (fear and neutral) and two levels of attention (towards and away from the faces). Participants had a

maximum of 4 seconds to make a response after which there was a jittered intertrial interval (jitter was created using an exponential function resulting in an ISI ranging from a minimum of 6 seconds to a maximum of 12 seconds). In total eight blocks were completed per subject leading to 160 trials. The task took approximately 20 minutes to complete. 50% of the facial stimuli were female, and for both male and female faces the bars were aligned on 50% of trial ensuring that the required responses were balanced between the right and left buttons and that the gender of the face was orthogonal to the alignment of the bars. The practice version of the task completed before training (see section 3.2.3) presented neutral faces not used in the main task. The practice task continued until participants had responded correctly to four sequential gender trials and four sequential bars trials.

### 3.2.5. Image acquisition

A blood oxygenation level dependent (BOLD) contrast signal was acquired using echo planar imaging (EPI) on a 3T Siemens TIM Trio System. A total of 45 slices were acquired using a voxel resolution of 3x3x3 $mm^3$, TR = 3 s, TE = 30 ms, Flip angle = $87^o$. The slice angle was set to $30^o$ (Deichmann, Gottfried et al. 2003). T1-weighted structural images were acquired for subject alignment using an MPRAGE sequence with the following parameters: Voxel resolution 1x1x1 $mm^3$, Echo TE = 4.7 ms, Repetition time TR = 2040 ms.

### 3.2.6. Data Analysis

#### *3.2.6.1. Questionniare Data*

Baseline measures were compared between groups using independent t-tests for continuous data and chi-square tests for categorical data. Change in anxious and depressive symptoms over time was assessed using a (2 x 2) mixed model analysis of variance (ANOVA) with the

between subject factor of training group and the within subject factor of time of assessment (i.e. before or after training).

*3.2.6.2. Image analysis*

Functional magnetic resonance imaging analysis was carried out using FEAT (FMRI Expert Analysis Tool) Version 5.91 (part of the Functional Magnetic Resonance Imaging of the Brain (FMRIB)'s Software Library, www.fmrib.ox.ac.uk/fsl). Data were pre-processed using FEAT default options: motion correction was applied using rigid body registration to the central volume; Gaussian spatial smoothing was applied with a full width half maximum of 5 mm; brain matter was segmented from non-brain using a mesh deformation approach; high pass temporal filtering was applied with a cut off of 100 s. Each of the four trial types of interest were modelled (NB fearful or neutral faces which could be either attended or unattended) as 0.2 second events coincident with the stimulus presentation which were then convolved with a canonical haemodynamic response function. In addition the temporal derivates of these regressors were included as were two nuisance regressors; one reflecting trials in which participants made errors, the second representing periods when the participants were reading instructions (a four second period at the start of each block). All regressors were included in a general linear model applied voxel wise to the pre-processed, pre-whitened imaging data.

As the primary hypothesis was that the training effect would be mediated by alteration of frontal function, the initial analysis identified regions in which activity was greatest when the task conflicted with the training participants had received. Activity in these regions should be highest in the "avoid-negative" group when they were attending towards the fearful or away from the neutral face (as their training had encouraged the opposite tendency). Activity in the

"attend-negative" group should mirror this as exactly the opposite trials would conflict with their training. This pattern of activity is captured by an interaction contrast (emotion x attention) which was constructed from the four basic regressors and was registered to the MNI 152 template using affine transformation. Individual contrast images were combined at the group level in a random effects analysis allowing comparison between groups. Results from this analysis were corrected for multiple comparisons across the whole brain again using the Feat default options. Specifically a cluster based correction (Worsley, Evans et al. 1992) with an initial threshold of Z=2.3 followed by correction over the whole brain using a significance level of $p < 0.05$ was used. All activations are reported using MNI coordinates.

#### *3.2.6.3. Region of interest definitions*

As described in the introduction and Chapter One, a network of brain regions including both amygdale and the rostral anterior cingulate are believed to be involved in the control of attention to emotional information. This prior work motivated a complementary region of interest (ROI) analysis. Regions of interest were therefore defined using the following anatomical criteria: the rACC ROI included all voxels which had a greater than 50% probability of lying in the ACC according to the Harvard-Oxford cortical atlas (included in the FSL package) and which had a *y* co-ordinate of > 30 (Bush, Luu et al. 2000). The amygdale ROIs were defined simply as those voxels with a > 50% probability of lying within the amygdale using the same atlas.

### 3.2.7. Connectivity analysis

#### *3.2.7.1. Overview*

Having identified potential control regions in the main analysis, a connectivity analysis was used to test whether these regions did indeed influence a neural measure of attention; activity

in face selective visual sensory association cortex (the fusiform face area; Kanwisher, McDermott et al. 1997). This was achieved using a targeted psychophysiological interaction (PPI) analysis (Friston, Buechel et al. 1997) to assesses the connectivity between control and sensory regions. Briefly, as the "attend-negative" training increases negative attentional bias (see Chapter Two), control regions should act to increase the sensory response to fearful faces in this group, whereas following the "avoid-negative" training the control regions should favour neutral faces (Vuilleumier 2005). It was therefore assessed whether the observed connectivity between control and sensory regions would produce this effect. The analysis resulted in four estimates of connectivity per participant; one each for the links between both left and right sided attentional control regions with left and right sided sensory target regions. These data were entered into a (2 x 2 x 2) split plot ANOVA with training group as a between subject factor and control region (left vs. right) and target region (left vs. right) as within subject factors.

*3.2.7.2. Psychophysiological interaction analysis, details*

The PPI analyses (Friston, Buechel et al. 1997) examined the pattern of correlation between activity in the attentional control regions (i.e. lPFC clusters) identified in the main analysis and face selective visual sensory cortex (see section 3.2.7.4 for description of the localiser analysis used to identify face selective cortex). Separate analyses were used for the right and left lPFC clusters. In both of these analyses three regressors were included: the timeseries of the peak voxel from the lateral frontal cluster of the participant's fMRI data (the "physiological" regressor), a regressor coding for whether each block of 20 trials was attended or unattended (the "psychological" regressor) and a final regressor reflecting the interaction of the two (the "interaction" regressor). The interaction regressor provides a measure of how the correlation between activity in the lPFC and sensory cortex differs

between the two task demands. The mean value of this interaction regressor was extracted from each participant's left and right fusiform face area masks, for both PPI analyses, resulting in four data points per participant.

*3.2.7.3. Predicted Connectivity*

The question addressed in the PPI analysis is whether the pattern of connectivity between the identified lPFC clusters and the visual cortex is consistent with the hypothesis that lPFC activity controls attentional resources. The specific prediction tested here is based on two pieces of information: a) the pattern of activity in the lPFC as confirmed in the main analysis and b) the predicted effect of training on sensory activity-- increased attention to particular stimuli (such as fearful faces) as a result of training should increase the activity associated with those stimuli (Vuilleumier 2005). In short, the analysis tests whether the connectivity is such that a) would lead to b). As illustrated in Figure 3.2, the predicted effect emerges if there is a relatively more positive correlation between the control and sensory regions during unattended as compared to attended blocks. It is important to note that while the main analysis demonstrates that frontal *activity* differs between groups, the *connectivity* should not differ. In other words, the lPFC is expected to influence activity of the visual sensory cortex in the same way in both groups, with the attentional differences arising because the lPFC reacts to different stimuli in the two groups.

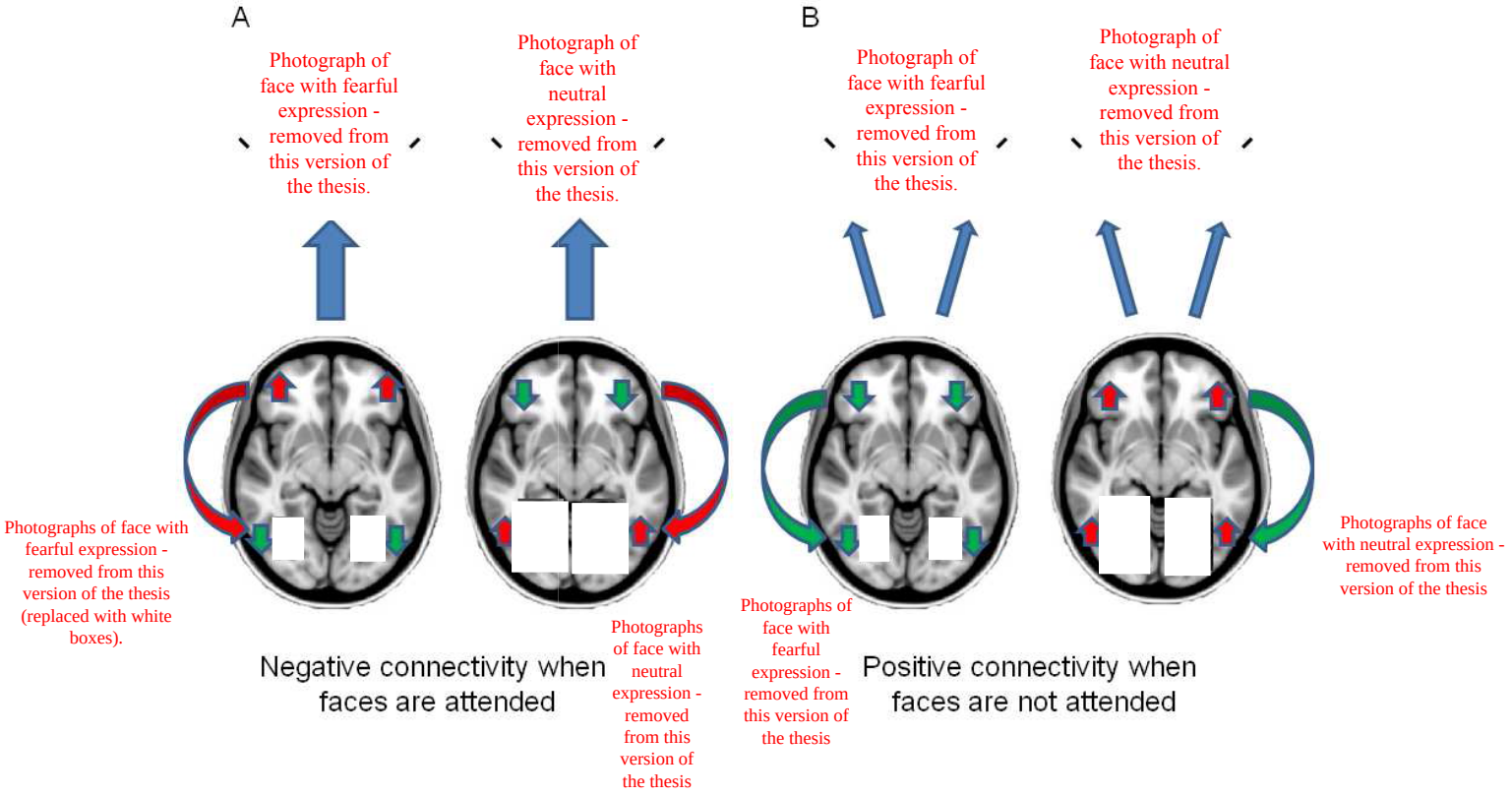



**Figure 3.2: Predicted connectivity during the different trials of the faces-bars task. The task includes blocks of trials in which the participant responds to the face (A) and others in which the participant responds to the bars (B). Each block includes fearful and neutral faces which are depicted separately. The blue arrows indicate the task demands of a specific trial (respond to the faces or to the bars). Neural activity is represented by arrows inside the brain with: 1) The upper arrows representing the response of the lPFC as demonstrated in the main analysis (the illustrated response is that seen in the avoid-negative group). 2) The lower arrows representing the predicted response of face selective sensory cortex (avoid-negative training should favour the sensory representation of neutral rather than fearful faces). The crucial issue in the analysis is determining whether the connectivity between these areas is consistent with the hypothesis that activity in 1) leads to 2). In the diagram "connectivity" is illustrated by the curved arrows outside the brain (red arrow "negative connectivity", green arrow "positive connectivity"). As can be seen, the expected pattern of sensory cortex activity emerges if the connectivity (i.e., correlation) between the regions is more positive in the bars trials (B) than the faces trials (A). The predicted pattern of connectivity in the attend-negative group is exactly the same; in this case the fearful faces will be favoured as the lPFC response is reversed. The PPI analysis tests for this predicted pattern of connectivity; specifically it assesses whether the correlation between the two areas is greater during the bars trials than during the faces trials.**

### *3.2.7.4. Localiser Procedure*

The PPI analysis examined connectivity between the frontal regions identified in the main analysis and face selective sensory cortex. A localiser procedure (Kanwisher, McDermott et al. 1997) was used to identify the face selective regions separately in each participant. This

was achieved by contrasting all trials in the faces task with all trials in a word based task which participants also completed in the same scanning session. The word task was of similar structure to the faces task (i.e., a centrally presented word was flanked by bars and presented for 200 ms) and allowed a separate localiser analysis to be performed for each participant. Specifically, the peak voxel for the contrast of all faces > all words within both the left and right fusiform cortices (determined as > 50% probability of lying within the fusiform cortex using the Harvard-Oxford cortical atlas provided with FSL) was identified. Individual fusiform masks of 8 mm diameter were then created, centred on these co-ordinates.

## 3.3. Results

### 3.3.1. Questionnaire measures

There were no significant differences between groups on any of the baseline measures, indicating that randomisation had been successful (see Table 3.1). Further, there were no between group differences on measures of anxiety or mood across training, indicating that the effects of the training cannot be attributed to a mood induction effect [all *p's* > 0.13].

| | **Avoid-negative Training** | **Attend-negative Training** | ***p*** |
|---|---|---|---|
| **Age** mean (sd) | 20.3 (1.6) | 20.5 (1.9) | 0.64 |
| **Female:Male** | 10:5 | 8:6 | 0.6[a] |
| **STAI-Trait** mean (sd) | 35.1 (5.7) | 33.2 (5.7) | 0.39 |
| **BDI** mean (sd) | 3.3 (2.3) | 3.0 (2.7) | 0.75 |

| | | | |
|---|---|---|---|
| **STAI-State pre training** mean (sd) | 31.5 (7.6) | 28.1 (6.4) | 0.21 |
| **STAI-State post training** mean (sd) | 33.8 (8.3) | 29 (7.9) | 0.12 |
| **VAS (sad –happy) pre training** mean (sd) | -5.4 (3) | -5.5 (2) | 0.9 |
| **VAS (sad –happy) post training** mean (sd) | -4.8 (3.6) | -4.6 (2.5) | 0.9 |

**Table 3.1**: **Demographic details for participants. BDI= Beck Depression Inventory. STAI-Trait and STAI-State are the trait and state subscales of the State-Trait Anxiety Inventory. VAS=Visual Analogue Scales. [a] = Chi-Square test was used to test null hypothesis of no difference between groups. For all other measures independent t-tests were used.**

### 3.3.2. Whole brain analysis

Consistent with the proposal that the attentional effects of training are mediated by alteration of frontal function, whole brain analysis comparing the emotion x attention contrast between groups revealed bilateral lPFC clusters, including dorsolateral [*x y z* = 36 54 16, *Z*-max = 3.22, *p*-corrected = 0.049] and ventrolateral PFC [*x y z* = 30 24 -2, *Z*-max = 3.4, *p*-corrected = < 0.0001] on the right and dorsolateral PFC [*x y z* = -30 54 10, *Z*-max = 3.27, *p*-corrected = 0.03] on the left (see Figure 3.3A). Importantly these clusters included voxels which lie within the regions of interest identified in previous studies of attentional control (Bishop, Duncan et al. 2004a). Additionally, clusters were found bilaterally in the striatum [Left; *x y z* = -20 6 0, *Z*-max = 3.55, *p*-corrected = 0.0002: Right; *x y z* = 28 8 4, *Z*-max = 3.85, *p*-corrected = < 0.0001].

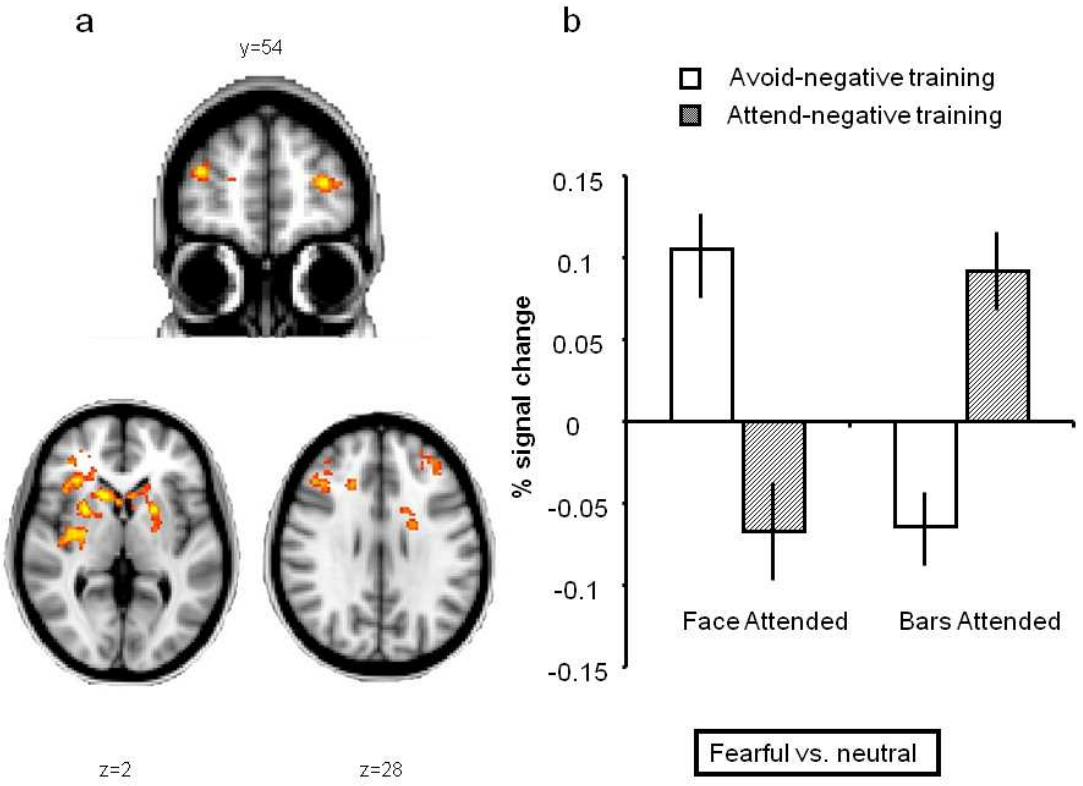



**Figure 3.3 Effect of attentional bias training on BOLD signal. (a) Whole brain, cluster corrected (*Z*-threshold = 2.3, $p < 0.05$) analysis demonstrating bilateral frontal and striatal regions in which activity corresponded to the effects of attentional bias training on the emotion x attention interaction. The activation map has been rendered onto the standard MNI brain. (b) The mean (SEM) percent signal change associated with the fear vs. neutral face contrast extracted from the right lPFC cluster (other clusters show an identical pattern). Estimates for the fearful face-neutral face contrast are displayed separately for trials in which participants had to attend to the location of the face (face attended) or to the location of the bars (bars attended). In all clusters activation is greatest when participants direct their attention contrary to their training; thus the avoid-negative training group (white bars), who have been trained to look away from negative and towards neutral stimuli, show increased activation when looking towards negative and away from neutral stimuli. The attend-negative training group (gray bars) show the opposite pattern of activation.**

As these clusters had been identified using an interaction contrast the nature of the interaction was then characterised by extracting individual estimates of the average signal change associated with fearful vs. neutral stimuli separately for trials in which attention was directed towards or away from the face. All clusters revealed an identical pattern of activation (Figure 3.3B); results from the extensive right lPFC cluster which spanned both dorsolateral and ventrolateral PFC are reported to illustrate this pattern. As predicted, across both training

groups and all experimental trials activity in these control regions is greatest when the direction of participants' attention conflicts with their training. Considering first the trials in which participants' attention is directed towards the faces (away from the bars): the attend-negative group has been trained to look towards negative stimuli and lPFC activation increases when they do the opposite; that is look towards the neutral faces [compared with fearful; $t$ (13) = 2.34, $p$ = 0.036]. In contrast the avoid-negative group, whose training induced the opposite tendency, show greater activation to the fearful faces [$t$ (14) = 5.25, $p$ < 0.001]. During the trials in which participants look away from the faces (towards the bars), the attend-negative group, who have been trained to look *away* from neutral stimuli, show greater activity when the face is fearful [compared with neutral; $t$ (13) = 4.04, $p$ = 0.001]. Again, the avoid-negative group display the opposite pattern of response with greater activation when neutral faces are to be avoided [$t$ (14) = 3.32, $p$ = 0.005]. Thus lPFC activity is determined by two factors; the behaviour of participants (as reflected in the type of information they are attending to) and the type of training undertaken. Across all trials and both training groups lPFC activity is greatest when the participants behave contrary to their training condition.

### 3.3.3. Connectivity Analysis

If, as predicted, the lPFC is mediating the attentional effects of training then activity in the identified frontal regions should influence activation of the face selective visual sensory cortex (Vuilleumier 2005). Specifically, in the avoid-negative group lPFC activity should favour the sensory representation of the neutral faces whereas in the attend-threat group the fearful faces should be favoured. This prediction was tested using a PPI analysis (see Figure 3.4); specifically the analysis assessed whether the observed pattern of connectivity between

the lateral frontal clusters and face selective visual sensory cortex would result in this effect. The expected pattern of connectivity was seen across both groups of participants [$F$ (1,27) = 2.45, $p$ = 0.045]. This was not modified by group, control region (left or right lPFC), target region (left or right sensory cortex) or any interaction of these factors [all $p > 0.12$]. These results are therefore consistent with the hypothesis that the information coded in lPFC activity is used in the control of attention to the facial stimuli. No further clusters of activation were identified in analyses of the PPI regressors across the whole brain and there were no significant interactions between lPFC and the amygdala when using a ROI approach.

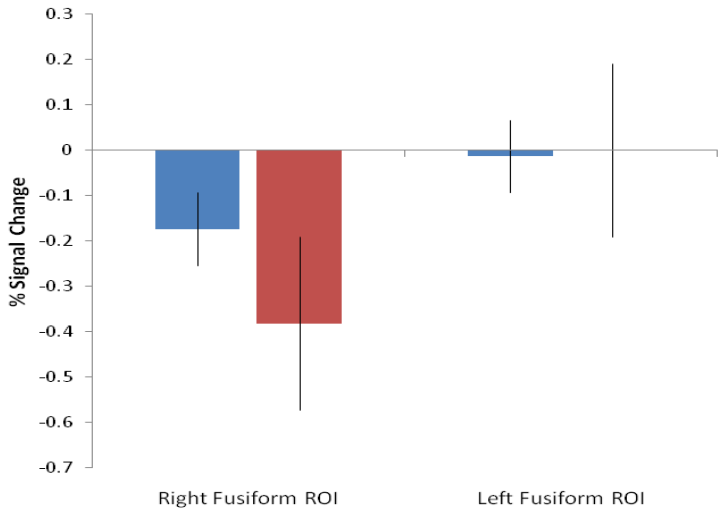



**Figure 3.4: Mean (SEM) estimate of the signal change associated with the PPI regressors in the left and right fusiform cortex. Connectivity between the left (blue bars) and right (red bars) lPFC and left and right fusiform cortex was assessed using PPI analyses. The analyses were coded such that a positive PPI indicates a more positive correlation during faces trials whereas a negative value indicates a more positive correlation during bars trials—the hypothesised pattern. Although the results are more prominent within the right fusiform ROI, statistical analysis revealed a general negative value across all measures and did not justify separate analyses of left and right fusiform ROIs.**

## 3.3.4. Region of interest analysis

### *3.3.4.1. Rostral ACC*

Although the rACC had also been predicted to be involved in mediating the effects of training, no activation was apparent on whole brain analysis. However, a small cluster was found in the rACC when a region of interest (ROI) approach was used (*x y z* = 0 32 8, *Z*-max = 2.91, *p*-corrected = 0.04). This cluster displayed an identical pattern of activation to that found in the lPFC and striatal clusters of the whole brain analysis.

*3.3.4.2. Amygdala*

Consistent with the prediction that attentional bias training would primarily be driven by alteration of frontal function no such effect was apparent in the amygdale, even when using an ROI analysis.

*3.3.4.3. Fusiform face area*

In the PPI analysis, the connectivity between lPFC and individually defined fusiform regions of interest was examined. When these ROIs were applied to the standard analysis there was evidence of greater activity during trials in which attention was directed towards the faces vs. those when attention was directed towards the bars [$F(1,27) = 6.7$, $p = 0.015$] as would be expected from face selective sensory cortex. However, there were no significant effects of training on gross activity [all $p$'s > 0.31].

### 3.3.5. Behavioural analysis

As intended, the groups did not differ on performance of the task in the scanner, with equivalent reaction times and error rates [all $p$'s > 0.10].

## 3.4. Discussion

The current study provides the first experimental evidence that attentional bias training can modify neural systems known to be involved in the control of attention to emotional stimuli

(Bishop, Duncan et al. 2004a; Etkin, Egner et al. 2006). Specifically, lPFC activity depended on the type of attentional bias training undertaken ("attend-negative" or "avoid-negative") and, across all participants, was greatest when the direction of participants' attention was contrary to their training. Connectivity between the identified lPFC clusters and face selective sensory cortex was consistent with that predicted by the behavioural effects of training and current models of selective attention (Desimone and Duncan 1995; Pessoa and Ungerleider 2005; Vuilleumier 2005; Bishop 2007). These results are in line with the prediction (DeRubeis, Siegle et al. 2008; Harmer 2008) that pharmacological (Sheline, Barch et al. 2001; Fu, Williams et al. 2004; Del-Ben, Deakin et al. 2005; Harmer, Mackay et al. 2006; Murphy, Norbury et al. 2009) and psychological interventions which alter attentional function are mechanistically distinct.

### 3.4.1. Does the lPFC mediate the attentional bias training effect?

While the main analysis reported above showed that attentional bias training modulated activity in the lPFC in an attentional task, it could not directly test whether these regions were actually involved in attentional control. It is conceivable, for example, that the training effect is encoded elsewhere in the brain and that the increased lPFC activity observed when the training rules were violated arise because behaving contrary to training is less practiced and thus more effortful; in essence, a form of tasking switching effect (Monsell 2003). By this interpretation altered lPFC activity results as a consequence of training rather than mediating its effect. This concern prompted the analysis of the pattern of connectivity between the identified lPFC regions and face selective visual sensory association cortex. In this analysis it was reasoned that, if the lPFC was controlling attention to the emotional faces as had been hypothesised, there should be evidence of a functional link between the control areas and the

visual sensory association cortex (Vuilleumier 2005) and, critically, this link should be such that the demonstrated activity in the lPFC would lead to the predicted activity in the sensory regions. The demonstrated pattern of connectivity is consistent with the hypothesis that the lPFC regions identified in the main analysis are indeed influencing attention. Clearly, however, a PPI analysis alone cannot prove that the lPFC controls activity in the fusiform cortex; the observed pattern of connectivity could equally well be produced by the fusiform controlling activity in the lPFC. However, the initial interpretation is in line with both the models of attentional control (Desimone and Duncan 1995; Pessoa and Ungerleider 2005; Vuilleumier 2005; Bishop 2007) and the more general understanding of the lPFC as providing a supervisory role in cognition (Duncan and Owen 2000).

Although the predicted pattern of connectivity between lPFC and sensory cortex was demonstrated, attentional bias training did not alter the gross activity of the face selective fusiform cortex (see section 3.3.4.3) which would have strengthened the interpretation of the results. Thus a single training session appears insufficient to individually demonstrate the effects of the intervention on all nodes of the attention circuit. Future studies using longer or more frequent training regimes may be able to show such an effect.

### 3.4.2. Involvement of other regions in the training effect

Against predictions, the rostral anterior cingulate cortex was not identified in the whole brain analysis. However, using a region of interest approach (see section 3.3.4.1) a small region of the rACC was found to display the same pattern of activity as the lPFC. Thus it seems likely that the lPFC regions identified in the main analysis are one node of a larger control circuit which incorporates the rACC. It may also include the striatum, because the whole brain analysis revealed bilateral striatal activity with a similar pattern of activity. Striatal

involvement in the training effect had not been predicted, so interpretation must be cautious; however the striatum is a component of a well described circuit which includes the lPFC (Seger 2006) and thus the striatal activity may reflect the efferent or afferent connections with the PFC.

### 3.4.3. Is the prefrontal cortex responding to conflict or to the violation of an expectation?

An interesting question is to consider whether the current results may offer any insight into the normal functioning of the prefrontal cortex. As reviewed in Chapter One, a number of alternative accounts of prefrontal function exist. A particularly influential account (Botvinick, Braver et al. 2001), which has informed the majority of previous investigations into the prefrontal control of emotional attention (Bishop, Duncan et al. 2004a; Etkin, Egner et al. 2006) describes the ACC and lPFC as detecting and responding to conflict. In the case of emotional stimuli conflict is often considered to arise as a consequence of the emotional dimension of the stimulus interfering with other, task related processing. In this light threatening stimuli are argued to produce greater emotional conflict than neutral stimuli (Bishop, Duncan et al. 2004a). However, this particular formulation of conflict cannot account for the current results as the fearful faces did not produce greater overall frontal activation than the neutral faces. In fact, the "conflict" seen in the current study arises when participants' behaviour violates the rules learned during the training procedure. While some authors have attempted to incorporate facets of learning into conflict models (Jones, Cho et al. 2002), an alternative possibility which may more naturally incorporate the role of learning is provided by expectancy based accounts of prefrontal function (Egner 2007; Rushworth, Mars et al. 2009). Specifically, the current results may be consistent with attentional bias

training generating a behavioural expectation (for example, the expectation that negative stimuli will be attended following attend-negative training) with the activity of the lPFC reflecting the violation of these expectancies; that is encoding a prediction error response. In the following chapter this hypothesis will be investigated by reanalysing the current data set in order to assess whether the expectancy and prediction error signals generated by a simple computational learning algorithm can predict the behavioural and neuroimaging measure of attention to negative stimuli.

### 3.4.4. Conclusion

In summary, the current study demonstrates that lPFC activity to emotional stimuli may be modified by a simple cognitive intervention known to alter attentional bias. This supports the proposal that modification of PFC function contributes to the effects of psychological interventions which target attentional processes (e.g. CBT) and suggests that such interventions are mechanistically distinct from pharmacological approaches. In the following chapters the consequences of these results will be explored using novel neuroimaging analyses (Chapter Four) and by directly testing the effects of combining pharmacological and cognitive manipulations of attentional function (Chapter Five).

# Chapter Four

# Threat Expectation and Surprise Predict Behavioural and Neural Measures of Attention to Emotional Stimuli

## 4.1. Introduction

In the preceding chapter the influence of attentional bias training on prefrontal activation was assessed using functional magnetic resonance imaging (fMRI). Activity of the lateral prefrontal cortex (lPFC) and striatum was found to be increased when participants behaved contrary to the training that they had previously completed. It was suggested that this pattern of activity may be understood as a form of prediction error signal; that is, a signal which quantifies the difference between a previously generated prediction and actual experience (Rescorla and Wagner 1972; Sutton and Barto 1998).

The notion of prediction and prediction error arise from the field of reinforcement learning (Sutton and Barto 1998), a computational framework which informs some of the most influential models of reward guided behaviour (Rescorla and Wagner 1972). In its simplest form, reinforcement learning suggests that stimulus-reward (or action-reward) associations are represented by the expected reward associated with the stimulus, also known as the stimulus' "value". This value is updated whenever the stimulus is encountered using the "Rescorla-Wagner rule" (Rescorla and Wagner 1972):

$$V_{t+1} = V_t + \lambda(R_t - V_t)$$

Where $V_t$ is the value of the stimulus at time t, $R_t$ is the reward delivered at time t and $\lambda$ is the learning rate (a parameter which determines the extent to which experience is used to update predictions). The final term in the equation, $(R_t - V_t)$ is the reward prediction error, the difference between the reward received and that expected. The absolute value of the prediction error (i.e. regardless of the sign of prediction error) is a measure of how "surprising" an event is (den Ouden, Friston et al. 2009).

The use of expectancy based models, such as that described above, has been stimulated by work in non-human primates which suggests that dopaminergic neurons originating in the ventral tegmental area encode reward prediction errors (Schultz, Dayan et al. 1997). More recently, evidence for reward expectancy (value) signals have been found in medial and orbital prefrontal cortex (Ostlund and Balleine 2007; Rudebeck, Behrens et al. 2008; Glascher, Hampton et al. 2009; Rushworth, Mars et al. 2009; Walton, Behrens et al. 2010). However, while the use of expectancy based models to describe reward guided behaviour has proved particularly fruitful, there is no particular reason why the validity of these models should be limited to reward processing. Indeed, the updating of predictions by prediction errors has been proposed as an organising principle for the entire cortex (Friston, Kilner et al. 2006). Consistent with the more general application of expectancy processes, surprise signals (i.e. unsigned prediction errors) have been described during non-reward based stimulus-stimulus learning (den Ouden, Friston et al. 2009) and both expectancy and surprise signals are apparent in the visual cortex during perceptual tasks (Summerfield and Koechlin 2008). Indeed, even higher order parameters, such as expectations about how others will behave, seem to be updated using similar processes (Behrens, Hunt et al. 2008; Hampton, Bossaerts et al. 2008). While such approaches have yet to be applied directly to attentional control, an

eye-tracking experiment indicates that participants tend to orientate attention towards the visually surprising aspects of a scene (Itti and Baldi 2009). Thus there is increasing evidence, across a range of paradigms, that both behaviour and neural activity may be successfully modelled by expectancy based processes.

The current chapter explores the novel proposal that the deployment of attention is influenced by expectations about the emotional content of stimuli. Specifically, it was hypothesised that attention would be directed towards a) stimuli that are expected to be threatening and b) stimuli that are surprising, regardless of their valence. In order to test these hypotheses a simple computational learning model (Behrens, Woolrich et al. 2007) was used to generate "threat-expectancy" and associated surprise signals; the ability of these signals to explain the trial-by-trial allocation of attentional resources to emotional stimuli, as measured by both reaction time (Posner and Petersen 1990) and blood oxygen level dependent (BOLD) signal in the visual cortex (Vuilleumier 2005), was then assessed. Having identified the presence of these model derived signals in both the reaction time and visual cortical data, the broader network of cortical regions presumed to be involved in generating the signals was explored. Lastly, as it has been suggested that the emotional disorders are associated with decreased control of attention to negative stimuli (Bishop 2007; De Raedt and Koster 2010), the association between the use of expectancy to control attention and individual differences in trait anxiety and depressive symptoms was examined.

## 4.2. Methods

### 4.2.1. Overview

The critical issue explored in the current chapter is whether an expectancy based model of attentional control can explain the trial-by-trial variability in attention to emotional information *throughout the course of the imaging task* (see section 4.2.2). This is achieved using a within subject measure of alteration of attentional function, which is in contrast to the analysis described in Chapter Three where the general effect of training across the whole imaging task was assessed using a between subject analysis. The current study therefore does not examine the effects of training. Consequently, all participants are analysed as a single group rather than being split by training group.

The current chapter describes a reanalysis of the study presented in Chapter Three. Details of participant characteristics, task structure and imaging parameters can be found in the methods section of that chapter. Relevant aspects of the procedure and analysis which were not described in Chapter Three are included below.

### 4.2.2. Imaging Task

As described in the previous chapter (Figure 3.1), the faces-bars task presented stimuli in blocks of 20 trials each. For a particular block the participant had to either report on the gender of the centrally presented face or whether the flanking bars were aligned. An aspect of the task which is central to the current analysis, but was not covered previously is that the frequency of fearful faces within a block was not constant (see Figure 4.1), that is blocks with relatively high and low proportions of fearful vs. neutral faces were presented (high-threat and low-threat blocks respectively). This modification to the task was based on the design described by Bishop and colleagues (Bishop, Duncan et al. 2004a) and allows assessment of whether the relative frequency of an emotional stimulus alters the deployment of attention. The sequence of trials within a block was such that the first three trials were always of the

same valence (fear in a high-threat block and neutral in a low-threat block), following this the majority (10 out of 17) of subsequent trials also displayed the dominant expression. Thus, at the start of a block, an expectation of a particular level of threat is developed and this expectation is largely maintained throughout the rest of the block. All blocks, regardless of the task completed during the block (i.e. identifying the gender of the face or alignment of the bars) could be high or low-threat. The specific sequence of trials within a block was selected from 8 predetermined pseudorandom schedules which conformed to the above rules. These schedules were randomly selected within the constraints that for a given participant there were two blocks each of the four possible block types (i.e. high-threat face-attended, low-threat face-attended, high-threat bars attended, low-threat bars attended) and that the task required of participants (gender of face or orientation of bars) alternated in successive blocks.

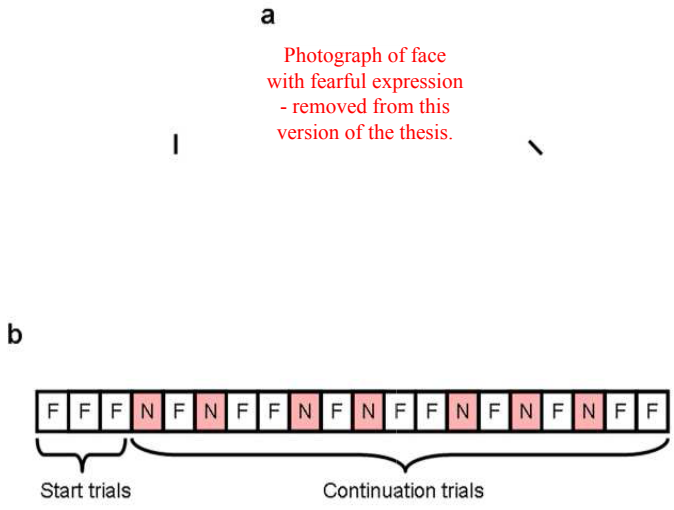



**Figure 4.1: a) sample trial from the task completed during the study. b) example sequence of trials from a single high-threat block of the task. Each trial displayed either a fearful (F) or neutral (N) face. The first three (start) trials were always of the same expression and in the remaining (continuation) trials the same expression was more common. All blocks were either high-threat as shown or low-threat, in which the neutral faces were more common.**

### 4.2.3. Computational Model

The computational model used was initially developed by Behrens and colleagues to explore the role of the prefrontal cortex in tracking the variability of reward (Behrens, Woolrich et al. 2007). The model achieves this by implementing a Bayesian version of the simple Rescorla-

Wagner rule described above. In the current study, the model was provided with information about whether a threatening (fearful) or neutral face was presented on each trial of the task. The model then used this information to update a single parameter, it's expectation that a threatening face would be presented on the following trial. The output from the model was thus in the form of a single vector of numbers between the values of 0 and 1 which represented the model's belief at any given trial that the following trial would display a threatening face (i.e. 0= the model is certain the following trial will be neutral and 1=certain it will be threatening). The model began each block with no bias about which face would be presented; that is, the initial expectation was 0.5. The expectation data provided by the model was used to derive the surprise signal which was defined as the absolute difference between the model's prediction about what face would be displayed and the face that was actually displayed (i.e. the unsigned prediction error)[2]. Model derived expectation and surprise regressors were created for each individual participant in the study, using the particular sequence of trials displayed to that participant (see Figure 4.2 for example of model derived information).

[2] In information theory "surprise" has a formal definition (MacKay 2003);

$$-\ln\left(p(data|model)\right)$$

That is, the negative logarithm of the conditional probability of the data given the model. In fact, this definition results in a regressor encoding essentially identical information to that derived using the simple unsigned prediction error formula described above (correlation between the unsigned prediction error and surprise regressors; $r = 0.99$, $p = 8 \times 10^{-78}$).

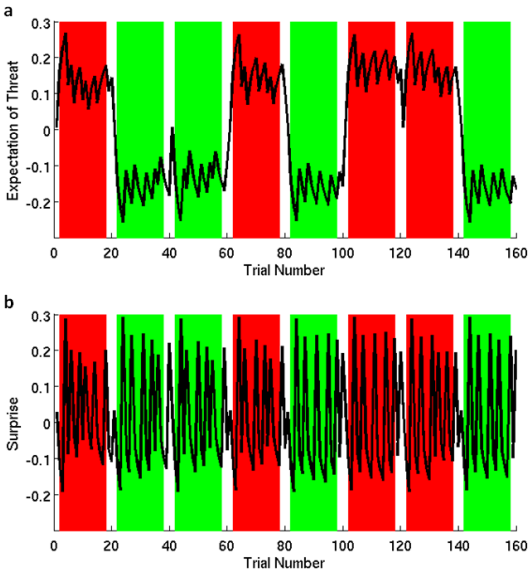



**Figure 4.2: Examples of one individual's expectation and surprise regressors derived from the computational model during the faces-bars task. All 160 trials in the faces-bars task are arranged along the x-axis, trials are presented in blocks of 20 in which fearful faces (high-threat blocks, represented by red areas) or neutral faces (low-threat blocks, represented by green areas) are more frequent. a) The model derived threat-expectation regressor. As can be seen the model accurately tracks the structure of the task—in high-threat (red) blocks it quickly learns to expect fear faces whereas in the low-threat (green) blocks it expects neutral faces. b) the surprise signal from the same trials. Surprise is defined as the absolute value of the difference between expectation and reality. NB All regressors have been demeaned as is required by both the behavioural and fMRI analyses (see sections 4.2.4.1 and 4.2.4.2)**

A key advantage of the Bayesian approach employed in the current model is that it requires no free parameters. Free parameters are used to tune a model to a particular individual's behaviour, for example, in the simple Rescorla-Wagner updating rule described above there is a single free parameter, the learning rate (λ). The value of this parameter is not specified by the environment or by the model, beyond the fact that it must lie between 0 and 1, and thus has to be chosen by the researcher. This is often done by selecting the value of the parameter which results in the model providing the best description of an individual's data (e.g. by using a maximum likelihood procedure; den Ouden, Friston et al. 2009). However, a concern with this approach is that it may overestimate the ability of the model to explain the data, as the fitting process entails a post hoc tailoring of the model to that particular series of data (and its associated noise). The model used in the current study infers a statistically

appropriate value for the learning rate from the data it is presented with and thus requires no parameter fitting (a full description of the mathematics by which these values are inferred can be found in; Behrens, Woolrich et al. 2007). Consequently the ability of the model to explain participants' behaviour or neural activity cannot be accounted for by tailoring of the model to a particular set of data.

### 4.2.4. Data analysis

Similar analysis strategies were employed for both the behavioural and neuroimaging data. Initially, a series of first level analyses were performed, one for each participant, in which a general linear model was used to assess the ability of the individualised expectancy and surprise regressors (as well as relevant nuisance regressors) to explain that participant's data. Following this, the results of the first level analyses were combined using a random effects analysis to assess whether the expectancy and surprise regressors were able to predict the behavioural and neural data, across the population of participants.

#### *4.2.4.1. Behavioural data*

In order to assess whether the model derived signals predicted a behavioural measure of attention, reaction time data from the faces-bars task was analysed. The logic of this analysis follows that employed in the Posner-like measures of selective spatial attention (MacLeod, Mathews et al. 1986; Posner and Petersen 1990). It was assumed that reaction time to identify the gender of the face would be reduced if attention was focussed on the face, whereas focussing attention on the face would increase reaction time for the bars task. Thus, increased attention towards the facial stimuli should be accompanied by increased reaction time on the bars task and decreased on the gender task (see Figure 4.1a for example trial from the task). As the intention was to assess the influence of trial-by-trial changes in expectation

and prediction error, the reaction times for individual trials were used as the dependent variable in a series of general linear models, with one analysis being performed for each participant in the study (see; Behrens, Hunt et al. 2008 for a similar analysis of model derived data). The positive skew of the reaction time distributions (Hogg, McKean et al. 2005) was reduced by removing extreme responses (less than 200 or greater than 1400 ms; mean data loss 7%) and log transforming the data before performing the analysis. The explanatory variables used in these first-level analyses coded for all the relevant model and task derived information; thus individual regressors for a) the emotion of the face (fear or neutral), b) the task being performed (gender identification or bars aligned), c) the model derived expectancy and d) surprise signals and e) whether the participant made an error on that trial were entered. Three further interaction regressors were also included; these coded for the interaction between the task performed and: f) the emotion of the face, g) the expectancy and h) the surprise signals. As explained above, the critical explanatory variables in these analyses were the interactions between the task performed and both the expectancy and surprise signals as these assess whether the model significantly predicts the deployment of attention towards the faces. In order to allow the inclusion of interaction regressors all of the basic explanatory variables (with the exception of the constant term) were demeaned before entry into the analysis (Friston, Buechel et al. 1997).

The individual regression analyses resulted in estimates of the regression coefficients for each explanatory variable, for each participant. These were then tested using one sample t-tests (two tailed) in a second level analysis to assess whether they significantly differed from zero and whether this difference was in the predicted direction. Analysis was performed using the statistical toolbox of Matlab (R2009a; The Mathworks Inc., Natick, MA).

*4.2.4.2. Imaging Analysis*

Functional magnetic resonance imaging analysis was carried out using FEAT (FMRI Expert Analysis Tool) Version 5.91 (part of the Functional Magnetic Resonance Imaging of the Brain (FMRIB)'s Software Library, www.fmrib.ox.ac.uk/fsl). As in the previous chapter, image pre-processing employed the default options for FEAT.

The first level model employed in the current analysis included individual regressors coding for each of the four trial types (fear attended, fear unattended, neutral attended, neutral unattended) as well as for trials where an error was made. Additionally, the demeaned regressors of the model derived expectancy and surprise signals were entered. These regressors were timelocked to the onset of the appropriate face stimuli; the events coded by the expectancy regressor had a duration of 4 seconds (as threat expectancy is the only information the model carries from trial to trial), the events coded by the surprise regressor had the same duration as the faces presented (0.2s; as a prediction error response is a transient updating of the expectancy). In this case, demeaning the model derived regressors ensures that neural activity attributable to the simple presentation of facial stimuli is captured by the basic task regressors (e.g. fear attended regressor) and that the model derived regressors can only account for activity over and above this.

A random effect analysis was employed at the group level to assess for regions in which the BOLD signal covaried with the model derived regressors across all participants. Trait anxiety (using the trait-STAI; Spielberger, Gorsuch et al. 1983) was added as a covariate in this analysis to assess whether varying levels of trait anxiety were associated with varying use of expectancy and surprise to control attention. A supplementary analysis was also performed using depressive symptoms (using the BDI; Beck, Ward et al. 1961) as the covariate. As per

Chapter Three, the FEAT default options for the group level analysis were used; cluster correction using a Z-threshold of 2.3 and correction over the whole brain using $p < 0.05$. For the identified clusters, MNI co-ordinates provide the location of the peak voxel.

## 4.3. Results

### 4.3.1. Behavioural Data

Both the model derived expectancy and surprise signals had the predicted effect on the deployment of attention as assessed using reaction times. Specifically, reaction times were increased in the bars relative to the gender task on trials in which the model expected the face to be threatening [expectation x task; $t(28) = 2.25$, $p = 0.03$] and on trials in which the face presented was surprising [surprise x task; $t(28) = 2.14$, $p = 0.04$]. As can be seen from Figure 4.3, participants were also generally slower on trials in which they made an error [$t(28) = 5.0$, $p < 0.001$] and when completing the bars rather than the faces task [$t(28) = 2.3$, $p = 0.03$]. The attentional effects of threat-expectancy and surprise did not seem to be driven specifically by either the faces or bars trials as, when considering the faces trials separately, both threat-expectancy and surprise were associated with a reduction in reaction time and, although there was a trend level effect for threat-expectancy, neither difference achieved statistical significance [threat-expectancy; $t(28) = 1.94$, $p = 0.06$. surprise; $t(28) = 1.1$, $p = 0.27$]. For the bars trials both threat-expectancy and surprise produced similar, non-significant increases in reaction time [$t(28) = 1.50$, $p = 0.14$ and $t(28) = 1.67$, $p = 0.10$]. There was no correlation between levels of trait anxiety and either the threat-expectation x task [$r(29) = -0.14$, $p = 0.5$] or surprise x task regressors [$r(29) = 0.06$, $p = 0.75$].

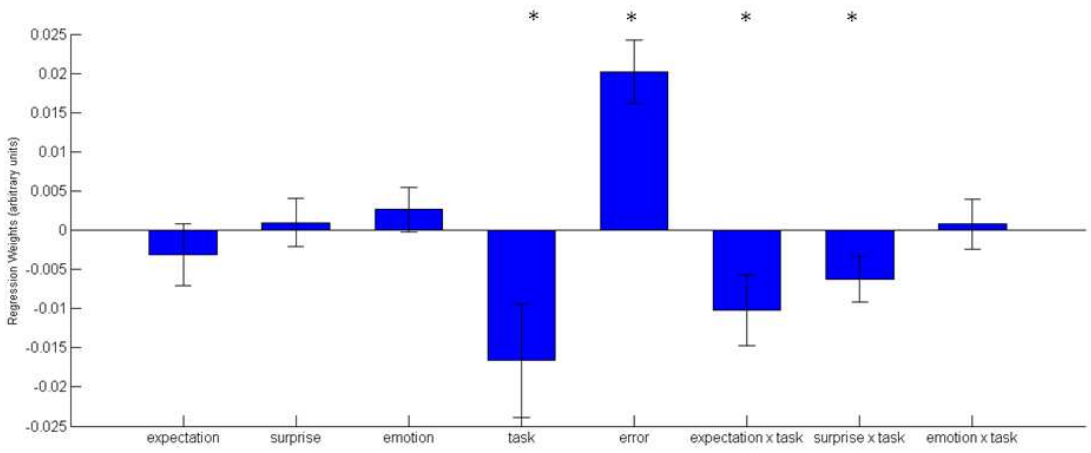



**Figure 4.3: Average (SEM) regression weights from the analysis of reaction times. Regression analyses were run on each individual's reaction time data using the explanatory variables listed along the x-axis (expectation, surprise, emotion, task, error, expectation x task, surprise x task, emotion x task). * = p<0.05 when explanatory variables were tested against zero using a one-sample t-test.**

## 4.3.2. Imaging data

### *4.3.2.1. Threat expectation*

Consistent with the behavioural analysis, a neuroimaging measure of the deployment of attention, activity in the visual cortex, was increased when threat-expectation was high (Figure 4.4a). A cluster in which activity covaried with the expectancy regressor was identified extending throughout the left ventral visual processing stream from occipital pole to temporal fusiform cortex [*x y z* = -8 -80 -2, *Z*-max = 3.93, *p*-corrected < 0.001]. Although the cluster extended to the lingual gyrus on the right, it did not include the right fusiform cortex. The whole brain analysis also identified a number of attentional control regions in which activity also covaried with the threat-expectation signal. These included bilateral orbitofrontal cortex [OFC; *x y z* = 12 46 -16, *Z*-max = 3.66, *p*-corrected = 0.001] and a left sided fronto-temporal cluster [*x y z* = -8 -80 -2, *Z*-max = 3.93, *p*-corrected < 0.001] which extended into the left amygdala. Additionally a posterior parietal cluster was identified,

although this was on the left hand side [*x y z* = -36 -50 28, *Z*-max = 3.91, *p*-corrected = 0.004].

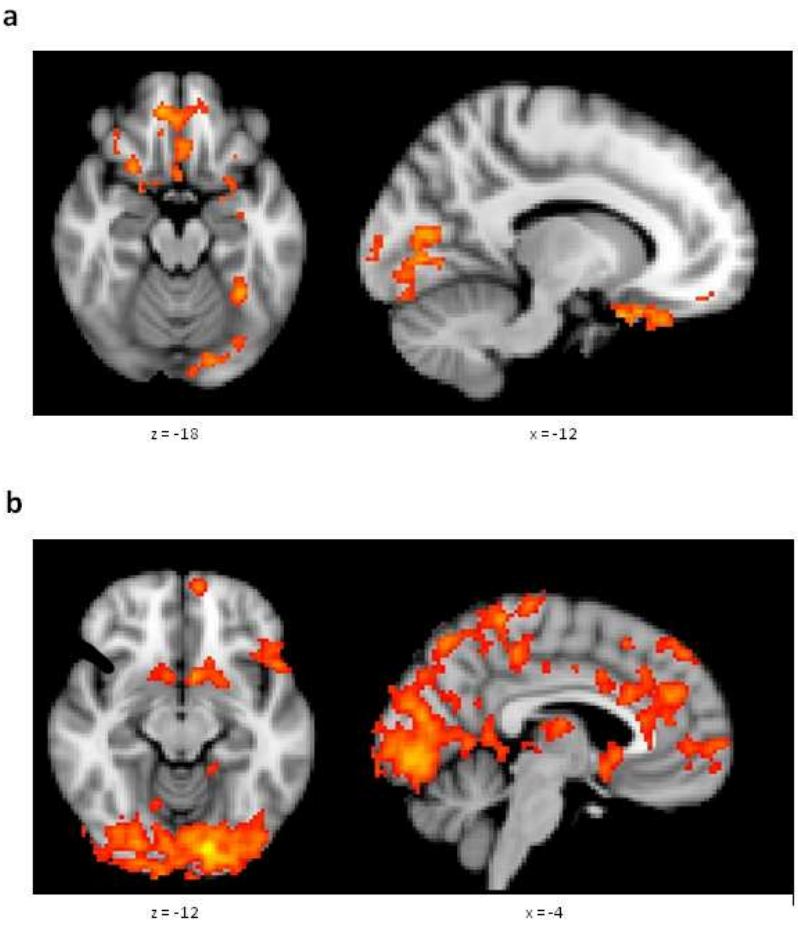



**Figure 4.4: Results of whole brain analyses showing activity associated with a) the threat-expectation and b) surprise regressors. a) threat-expectation was found to covary with activity in the left ventral visual stream and with attentional control regions including the OFC and left amygdala. b) surprise was associated with activation of the visual cortex and a network of frontal regions including the ACC, dlPFC and ventral striatum (not shown). Average activations maps across all participants have been rendered over the standard MNI brain.**

*4.3.2.2. Surprise*

An extensive network of regions was found to covary with the surprise regressor (Figure 4.4b). This included large volumes of both the left and right visual areas including both occipital fusiform cortices [*x y z* = -14 -88 -12, *Z*-max = 4.46, *p*-corrected < 0.001], an

extensive frontal midline cluster incorporating both dorsal and rostral ACC [*x y z* = -2 20 28, *Z*-max = 3.63, *p*-corrected < 0.001], left sided dlPFC [*x y z* = -16 66 2, *Z*-max = 4.01, *p*-corrected < 0.001] and vlPFC [*x y z* = -48 16 -16, *Z*-max = 3.82, *p*-corrected = 0.02] and bilateral striatum [*x y z* = -6 8 -14, *Z*-max = 3.77, *p*-corrected = 0.03]. These regions overlapped the left dlPFC and bilateral striatal clusters identified in the previous chapter as being activated when participants behaved contrary to their training contingencies (see Chapter Three).

*4.3.2.3. Are the visual cortex activations face specific?*

Although the visual cortical activation associated with threat-expectation incorporated the left temporal-occipital fusiform cortex, the influence of both the threat-expectancy and surprise signals was much more marked in posterior regions of the visual cortex which would not be expected to be face specific. Consistent with this, when analysis was limited to the individually identified fusiform face area (FFA) regions of interest (ROI) described in Chapter Three the effects of threat-expectancy and surprise, while in the right direction, did not reach significance [$F(1,28) = 2.86$, $p = 0.10$]. However, as the face stimuli are much more complex than the bars, attending to the face stimuli would be predicted to increase activation even in lower order visual areas. An alternative analysis suggests that the observed increase in lower order visual cortical activity found with both threat-expectancy and surprise is indeed associated with the processing of the faces: As described in the previous chapter, the demands of the task, to identify the gender of the face or report on the alignment of the bars, cause attention to be directed towards or away from the faces (consistent with this activity in the FFA ROIs was found to be greater during the faces task than the bars task). When comparing the effect of these task instruction across the whole brain (using the same multiple correction parameters as described above) a single cluster was found to be more active during

the faces task than the bars task, this cluster is located in the occipital pole [*x y z* = -18 -104 0, *Z*-max = 3.38, *p*-corrected = 0.02] and, in a conjunction analysis, overlaps with both the threat-expectation and surprise related visual cortical clusters. Thus it seems likely that the increased visual cortical activity identified in the above analysis is indeed associated with increased attention to the faces.

#### *4.3.2.4. Is trait anxiety associated with the use of expectation and surprise to control attention?*

A number of regions displayed a negative correlation between trait anxiety and utilisation of the threat-expectancy signal, suggesting that participants with higher general levels of anxiety were less likely to use threat-expectancy to guide attention. These included large areas of the right posterior parietal and lateral occipital cortex [*x y z* = 32 -68 52, *Z*-max = 3.79, *p*-corrected < 0.01], right side dlPFC [*x y z* = 50 8 18, *Z*-max = 3.66, *p*-corrected < 0.01], right [*x y z* = 32 16 -10, *Z*-max = 3.84, *p*-corrected < 0.01] and left vlPFC [*x y z* = -34 24 -6, *Z*-max = 3.78, *p*-corrected < 0.01], medial PFC including the dorsal and rostral ACC [*x y z* = 4 42 28, *Z*-max = 4.15, *p*-corrected < 0.01] and the posterior cingulate/precuneous [*x y z* = 8 -34 26, *Z*-max = 3.59, *p*-corrected = 0.02]. These regions overlapped with those identified as responding to threat-expectancy generally in the right (32 26 -10) and left vlPFC (-16 14 -20), left temporal lobe abutting the amygdala (-32 -8 -18) and precuneous (10 -56 16). A similar negative correlation was found when analysis was limited to the visual cortical cluster, identified in the conjunction analysis reported above (section 4.3.2.3) as being activated by both the expectation and surprise regressors [$r(29) = -0.37$, $p = 0.05$]. No regions displayed a positive correlation with trait anxiety.

Consistent with the decreased use of expectancy to guide attention in high levels of anxiety, trait anxiety was found to correlate positively with surprise in a single cluster in the left superior temporal gyrus [*x y z* = -52 -44 4, *Z*-max = 3.80, *p*-corrected = 0.01] with no regions displaying a negative correlation. However, interpretation of this finding is limited by the fact that this region was not identified in the initial analysis as responding to either expectancy or surprise.

*4.3.2.5. Can the effects of trait anxiety be differentiated from the effects of depressive symptoms?*

While the above analysis demonstrates that the correlation between neural activity in attentional control areas and threat expectation is influenced by trait anxiety, it does not provide evidence that this effect is specific to symptoms of anxiety. As reviewed in Chapter One there is increasing evidence that depression is also associated with abnormalities in the deployment of attention and therefore it is reasonable to ask whether a similar effect is seen for symptoms of depression. In fact, in the current, non-clinical sample trait anxiety, as measured by trait-STAI, and depressive symptoms, as measured by the BDI, were strongly positively correlated with each other [$r(29) = 0.53$, $p = 0.003$]. This is consistent with the proposal that such scales tend to provide general measures of negative affectivity (Teasdale and Dent 1987) and that fractionating the effects of depression and anxiety will be challenging in the current population. Consistent with this, using BDI as a covariate in the group level analysis produced a less extensive, but similar, pattern of activation to that found with trait-STAI. Specifically, when analysing threat expectancy clusters were identified in the right posterior parietal/superior temporal lobe [*x y z* = 62 -26 28, *Z*-max = 3.62, *p*-corrected < 0.01], the right lateral occipital lobe [*x y z* = 20 -70 54, *Z*-max = 3.3, *p*-corrected < 0.01] and the rostral anterior cingulate cortex [*x y z* = -12 26 12, *Z*-max = 3.22, *p*-corrected = 0.01]. No

regions were identified when surprise was analysed. Thus, the degree to which threat expectation predicts neural activity in the current study appears to be a function of negative affectivity rather than specifically anxious symptoms.

## 4.4. Discussion

The current study demonstrates that a simple expectancy based model is able to account for a significant amount of variance in both the behavioural and neural signatures of attention to emotional information. Specifically, participants' directed their attention towards the facial stimuli when the computational model expected the stimuli to be threatening or when it found the emotional content of the stimuli surprising. The predictive validity demonstrated by the computational model in the current study suggests that similar expectancy based calculations may be instantiated by neural systems in order to influence the deployment of attention to threat. The degree to which individuals employ threat expectancy to control attention may vary as a function of negative affectivity as increased levels of self-reported trait anxiety and depressive symptoms was associated with a decreased correlation between the model derived threat-expectancy signal and neural activity in a number of attentional control regions. However, this relationship was not observed in the behavioural data. Lastly, the current results cannot be accounted for by task or stimulus related confounding factors such as the valence of the presented stimuli as both the behavioural and neuroimaging analyses performed were careful to control for such factors.

### 4.4.1. The generation of threat expectation in the brain

As well as demonstrating the predicted influence of threat-expectation and surprise in the visual cortex the current study was able to assess other regions of the brain which may be involved in generating and tracking such expectancies. The threat expectation signal itself

was found in bilateral OFC and the amygdala on the left. Interestingly, it has been suggested that both the OFC (O'Doherty, Kringelbach et al. 2001; Rolls 2004; Kringelbach 2005; Rudebeck, Behrens et al. 2008) and the amygdala (Murray 2007; Adolphs 2010) code for the value of stimuli in reward guided paradigms. Value is defined as the reward-expectation associated with a stimulus and, although the current study does not assess value, threat-expectation fulfils a similar role in the computational model as value does in reward learning. Thus the OFC and amygdala may perform a similar role in generating and tracking this different form of expectancy. Such a conclusion would be consistent with the observations that both these areas respond to salient negative as well as positive events (Kringelbach and Rolls 2004; Adolphs 2010).

The surprise signal was associated with activation across large areas of the midline of the brain including both the rostral and dorsal ACC, which have previously been described as reacting to "emotional conflict" (Bishop, Duncan et al. 2004a; Etkin, Egner et al. 2006). Importantly, the regions identified as responding to surprise in the current study include voxels which were part of the left sided dlPFC and striatal activations identified in the previous chapter, indicating that these regions may respond to violations of predictions regardless of whether the predictions relate to action-stimulus couplings, such as those encouraged by training, or recent stimulus history as in the current study. Consistent with this, prediction error activity has previously been reported in the (right) dlPFC and striatum during associative learning tasks (Corlett, Aitken et al. 2004; O'Doherty, Dayan et al. 2004).

### 4.4.2. Alternative models

The purpose of the current study was to test the novel hypothesis that expectancy based processes influence attentional control, rather than to evaluate the relative merits of a range of

possible models. It is therefore conceivable that alternative computational models would provide a more compelling description of attentional function. However, regardless of the specific structure of the model used, the demonstration that expectancy processes have a role in the control of attention provides useful insights into the existing literature. For example, Bishop and colleagues (Bishop, Duncan et al. 2004a) reported increased rACC activity in response to infrequent threat related distractors which they interpreted as reflecting the increased conflict caused by infrequent, but salient stimuli. The current study identified a network of regions, including voxels in the rACC mask used by Bishop et al., that responded to the violations of threat-expectations (i.e. the surprise signal, see section 4.3.2.2), and thus suggests that both results may be understood in terms of expectancy based processes. Similarly, Bishop et al. reported a reduced rACC and lPFC response to threatening stimuli in participants with high levels of trait anxiety, which they interpreted as demonstrating that anxious participants have particular difficulties in using previous threatening stimuli to trigger attentional control. In the current study we find a similar inverse correlation between anxiety and the use of expectancies but, critically, do so using an explicit model of how the expectancy was generated.

### 4.4.3. Does negative affectivity modulate attentional control?

As described above, a negative correlation was found between individual self reports of trait anxiety (and depressive symptoms) and the expectancy signal in a range of attentional control and visual cortical structures. There was also evidence of a positive correlation between trait anxiety and the surprise signal, although this was in an area which had not been identified in the main analysis (left superior temporal gyrus). While this may be consistent with those models which suggest decreased attentional control as a causal feature of the emotional

disorders (Bishop 2007; De Raedt and Koster 2010) there are some caveats to this interpretation which should be acknowledged. Firstly, in the current study there was no evidence that negative affectivity interacted with threat-expectancy when assessed using the behavioural measures of attention, which would have strengthened the interpretation of the expectancy signal effect. Secondly, the negative correlation between negative affectivity and the expectancy signal strictly only indicates that the computational model is less able to describe neural activity in participants with high negative affectivity, it does not tell us why this is the case. The possibility thus remains that participants with high negative affectivity *are* using expectancy information but that they are not updating the expectancies in a manner similar to the computational model used in the current chapter. Ideally this would be tested by developing a model of threat-expectation in anxious or depressed individuals and demonstrating that the new model better explains neural activity in these groups. Thus while the effect of negative emotional symptoms observed in the current study is consistent with certain models of attentional control in anxiety and depression, interpretation of these results must remain cautious.

### 4.4.4. Limitations to the interpretation of the model

The current study has described the influence of threat-expectation and surprise on the deployment of attention. It was possible to discriminate these signals as they vary independently from each other, allowing the construction of general linear models which can reliably partition variance between the signals. However, not all model derived signals may be discriminated. In particular, under the demeaning procedures required by the current analyses (see sections 4.2.4.1 and 4.2.4.2), the threat-expectation signal is strongly anti-correlated with "signed" prediction errors derived from the model (den Ouden, Friston et al.

2009). As a consequence it is not possible to be certain whether the threat-expectation regions identified in the current study are involved in storing expectancies between trials, updating them using the signed prediction error information or both of these processes together. Thus, while the current analysis identifies the regions which are generally involved in the process of updating expectancies using prediction errors, it cannot delineate the specific roles played by these areas in this process. A similar limitation prevents the current study from discriminating the effects of anxiety from those of depression (see sections 4.3.2.4 and 4.3.2.5); that is, the high correlation between the measures of depressive and anxious symptoms in the population of participants recruited for the study results in an inability to separate the effects of these two symptom types. Future studies may overcome these problems by using alternative experimental designs in which threat expectation and the signed prediction error are not highly correlated or which specifically recruit populations in which symptoms of anxiety do not covary with symptoms of depression. In fact, the design discussed in the following section (4.4.5), which assesses the direct effects of bias training, would be well suited to the first of these challenges.

### 4.4.5. Modelling the effects of attentional training

In the current study participants completed a single session of attentional bias training before entering the scanner which precluded using the computational model to capture the changes in attention induced by training. However, there is no a priori reason why the direct effects of training could not be captured by a learning model; indeed doing so may provide valuable insights into the processes involved in training. Future studies in which imaging data is acquired while participants complete the training task would allow assessment of whether the alterations induced by training may be usefully captured by such learning models.

### 4.4.6. Conclusion

In the current chapter, the internal parameters of a computational model of threat-expectancy were used to predict the behaviour and neural activity of non-clinical participants who were presented with emotional stimuli. The results of the study support one of the implications raised by Chapter Three; that expectancy based mechanisms are involved in the control of attention to emotional information. In the following chapters some of the practical implications of Chapter Three are investigated, firstly by assessing the interaction between attentional training and antidepressant medication (Chapter Five) and then by assessing and modifying attentional function in a group of patients with bipolar disorder (Chapter Six).

# Chapter Five

# Exploring the Interaction between Pharmacological and Cognitive Strategies for Altering Attentional Bias

## 5.1. Introduction

In Chapter One, the evidence that pharmacological and psychological interventions may be used to alter attentional bias in both clinical and non-clinical cohorts was reviewed. In the review it was suggested that a) modifications of attentional function may be one route by which the interventions exert their beneficial clinical effects and b) behavioural and neuroimaging studies indicate that these two modes of intervention may target distinct attentional control systems; a hypothesis which was supported by the neuroimaging data reported in Chapter Three. The current chapter continues the investigation into the mechanisms underlying psychological and pharmacological alterations of attentional function by assessing the interaction between the two when they are administered concurrently. Specifically, a factorial experimental design (see Figure 5.1) was employed in which non-clinical participants were administered seven days of the antidepressant medication citalopram or a placebo while also completing seven days of positive attentional bias training or a neutral training task. The effects of the interventions were then assessed using a range of behavioural measures of cognitive bias, self reported mood and anxiety and finally by

measuring the effect of external "emotional challenges" on depressive and anxious symptoms.

While no previous studies have examined the combination of attentional bias training and antidepressant medication, a number have characterised the cognitive effects of the interventions administered separately. The selective serotonin reuptake inhibitor (SSRI) citalopram, when administered over seven days, has been demonstrated to induce a range of relatively positive cognitive biases in non-clinical populations including reduced attentional bias to threatening faces (Murphy, Yiend et al. 2009), decreased recognition of fearful and angry facial expressions and increased memory for positive relative to negative words (Harmer, Shelley et al. 2004). The choice of seven days administration of the antidepressant in such studies is based on clinical evidence which suggests that treatment response occurs after longer durations (Blier and de Montigny 1994; Duman 2004; NICE 2009); the finding of altered cognitive function after only seven days is therefore argued to be consistent with these cognitive effects mediating the later treatment effects (Harmer 2008). Consistent with this, in depressed patients, the increased recognition of happy faces induced by two weeks of citalopram was found to predict treatment response at six weeks (Tranter, Bell et al. 2009). Thus citalopram has been found to induce relatively positive biases across a range of cognitive functions in both clinical and non-clinical populations, with these cognitive effects occurring quickly after initiation of the antidepressant and possibly mediating the later treatment effects.

A number of studies have demonstrated that "positive" attentional bias training regimes, when compared to neutral versions of the task[3], produce the intended positive attentional bias in both clinical and non-clinical populations (see Chapter One for review). However, there has been very little work examining the effects of attentional bias training on other cognitive processes such as memory or interpretation. While this focus on the alteration of attention is unsurprising, given that it is the explicit goal of the training regimes, attention plays a fundamental role in a wide range of other cognitive processes, including memory and interpretation, suggesting that interventions which alter attentional bias may well produce changes in the function of these systems. Consequently, the current study aimed to extend the assessment of the cognitive effects of attentional bias training to include measures of the interpretation and memory of emotional information which have been widely used in the psychopharmacological literature.

In contrast to the work described above, the psychopharmacological literature is less developed than that for attentional bias training in assessing the impact of the interventions on emotional reactivity. Emotional reactivity is generally assessed (e.g. MacLeod, Rutherford et al. 2002) using "emotional challenge tasks" in which the change in subjective mood in response to a "challenge" (for example attempting to complete insoluble anagrams under time pressure) is measured. Conceptually challenge tasks are well suited to studies investigating the alteration of cognitive bias as it is argued (Mathews and MacLeod 2002; Mathews and MacLeod 2005; MacLeod, Koster et al. 2009) that the effect of such alterations is not to directly induce a positive mood but rather to alter underlying cognitive function, with

[3] Positive attentional bias training includes any regime that induces a relatively positive bias; either by reducing bias towards negative information or increasing bias towards positive information. Neutral training is designed not to alter attentional bias and therefore may be conceptualised as providing a neutral control to the "active" intervention of positive training (c.f. Chapters Two and Three).

changes in mood or anxiety only occurring when subsequent, emotionally salient, information is experienced through the filter of the bias (Harmer 2008). Inducing a positive bias with either pharmacological or cognitive interventions should not, therefore, lead directly to improved mood or anxiety following the intervention. Rather, it is hypothesised, the positive bias will confer a greater resilience to the effects of external negative events; in the laboratory this is expected to manifest as a reduction of the negative affect provoked by placing the participant in a stressful or depressogenic situation. While a number of studies have demonstrated that attentional bias training alters the emotional impact of such challenge tasks (MacLeod, Rutherford et al. 2002; Krebs, Hirsch et al. 2010; Reese, McNally et al. 2010), the effects of pharmacological interventions have not been examined using such paradigms. A second aim of the current study was therefore to assess the impact of both attentional bias training and antidepressant medication on the same emotional challenge tasks.

From first principles, pharmacological and cognitive interventions may interact in three distinct ways: Firstly, they may interact positively, producing a greater effect when combined than the sum of each administered separately, secondly they may interfere with each other such that combined administration is less effective than the sum of the individual effects and lastly, they may be independent with the effect of combined administration being equivalent to the sum of both interventions administered separately. As no previous studies have examined the combination of attentional bias training and antidepressant medication, it is difficult to form strong hypotheses about the effects of combination administration. However, indirect evidence may be gleaned from the clinical efficacy literature in which it has been shown that both more complex cognitive interventions, such as cognitive behavioural therapy

(CBT), and antidepressant medications are individually effective in the treatment of anxiety and depression (NICE 2007; NICE 2009), whereas the results of studies assessing combined psychological and pharmacological treatments have been mixed with meta-analyses revealing only a small advantage for depression (Pampallona, Bollini et al. 2004; Cuijpers, van Straten et al. 2009) and none for anxiety (Foa, Franklin et al. 2002). These rather disappointing results from the clinical literature are difficult to reconcile with a positive treatment interaction and suggest either independent or interfering treatment effects are likely to occur, at least when complex cognitive treatments such as CBT are combined with antidepressant medication. While attentional bias training is obviously a much more restricted intervention than CBT, it has been suggested that alteration of cognitive bias may mediate the treatment effects of both CBT and pharmacological treatments for the emotional disorders (see Chapter One and; Harmer 2008). If true, this would suggest that the cognitive effects of the combined interventions investigated in the current study should show the same interaction suggested by the clinical results; either a negative interaction or independence.

The predictions for the current study reflect both the pre-clinical and clinical research evidence described above: both active interventions (positive attentional bias training and antidepressant medication) are predicted, when compared to the neutral interventions (neutral attentional bias training and placebo) to induce relatively positive biases in cognition and to reduce negative emotional reactivity to external challenge, however the combination of the two interventions is predicted to produce either no additional benefit or an interference effect. In keeping with previous studies using non-clinical volunteers (MacLeod, Rutherford et al. 2002; Harmer, Shelley et al. 2004; Browning, Reid et al. 2007), neither intervention is predicted to alter baseline measures of subjective mood or anxiety.

## 5.2. Methods

### 5.2.1. Summary of methodological changes in the current study

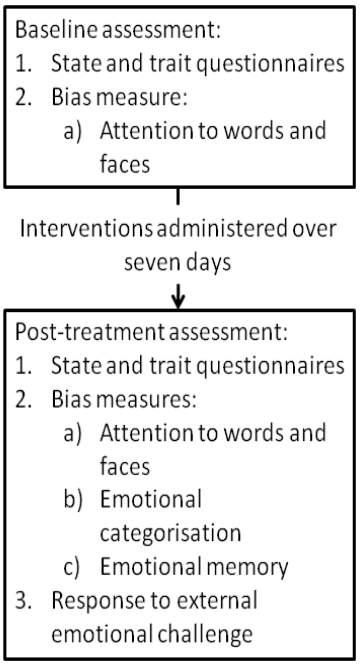
As discussed in the introduction, the major methodological alteration employed in the current study is the inclusion of an antidepressant treatment group. However, based on the findings reported in earlier chapters, a number of further modifications were made to the training and assessment tasks used. Firstly, participants were required to complete multiple training sessions over seven days, rather than a single session, in order to produce a more robust training effect (see Chapter Two). Secondly, with a view to increasing the generalisation of the training effect, the training task was modified (see section 2.5.2) so that a broader range of stimuli (i.e. positive, negative and neutral faces) were presented for both longer and shorter durations. Thirdly, positive training was compared to a "neutral" rather than negative (see Chapters Two and Three) task allowing comparison between active and placebo-like training interventions. Fourthly, as discussed above, a broader range of assessment tasks, drawn from previous studies of the effects of antidepressant medication (Harmer, Hill et al. 2003; Harmer, Shelley et al. 2004; Browning, Reid et al. 2007) and the attentional training literature (MacLeod, Rutherford et al. 2002; Krebs, Hirsch et al. 2010; Reese, McNally et al. 2010), were employed to assess for the generalisation of the induced emotional biases across a range of cognitive functions and to measures of emotional reactivity. The following sections provide a detailed description of the methodology employed in the current study.

### 5.2.2. Overview

The structure of the study is displayed in Figure 5.1a. Participants were randomly assigned to receive one of four intervention regimes (see Figure 5.1b); placebo with neutral training (placebo_neutral), placebo with positive training (placebo_positive), citalopram with neutral

training (citalopram_neutral) or citalopram with positive training (citalopram_positive). Each of the four regimes was administered over seven days. Participants attended assessment sessions immediately before starting and following completion of the treatment regimes. During the initial assessment, baseline measures of attentional bias were recorded as were trait and state measures of mood and anxiety symptoms. During the post-treatment session these same measures were repeated and were supplemented by assessment of cognitive bias in the categorisation of emotional stimuli (both words and faces) and emotional memory. The post-treatment session also included an assessment of participants' emotional reactivity to external emotional challenge. The extended measures of cognitive bias and the challenge tasks were not administered in the initial session as repeated use of these tasks in previous studies has resulted in marked learning effects which reduced their sensitivity to detect the influence of experimental interventions (unpublished observation).

a




b

| | Citalopram | Placebo |
|---|---|---|
| Positive Training | Citalopram_positive (n=16) | Placebo_positive (n=15) |
| Neutral Training | Citalopram_neutral (n=15) | Placebo_neutral (n=16) |

**Figure 5.1: a) Structure of study. Participants completed an assessment session both before and after undertaking the interventions. The specific measures employed in each assessment visit are listed. b) Randomisation to the four intervention groups. Participants were randomised to either active or control versions of both the pharmacological (columns; active=citalopram, control=placebo) and cognitive (rows; active=positive training, control=neutral training) interventions. Each of the four combinations of interventions were administered for seven days. The numbers of participants who completed the study in each of the four groups are shown.**

### 5.2.3. Participants

A total of 64, fluent English speaking participants who were judged to be healthy on the basis of a medical screen and brief physical examination were recruited. Participants were also screened to exclude current or previous axis I psychiatric disorder or alcohol/substance misuse using the Structured Clinical Interview for the DSM-IV (Spitzer, Williams et al. 2002). Other exclusion criteria included; taking any psychoactive medication, any significant neurological condition, familiarity with any of the tasks or stimuli used in the study or any contraindication to taking selective serotonergic reuptake inhibitor (SSRI) medication. All participants provided written informed consent to the study which had been approved by a Local NHS Research Ethics Committee. Two participants in the citalopram group experienced side effects (1 nausea, 1 insomnia) which prevented completion of the study and resulted in data being available for 62 participants.

### 5.2.4. Questionnaire measures

Participants completed questionnaire assessments of depressive (Beck Depression Inventory [BDI]; Beck, Ward et al. 1961) and anxious symptoms (Trait subscale of the State-Trait Anxiety Inventory [trait-STAI]; Spielberger, Gorsuch et al. 1983) as well as state measures of anxiety (state-STAI) and mood (Positive and Negative Affect Scale [PANAS]; Watson, Clark et al. 1988) both before and after treatment. In addition participants completed pre and post-treatment visual analogue scale (VAS) measures of a range of side effects associated with citalopram. These assessed participants' current levels of; nausea, dizziness, dry mouth, headache, alertness, agitation and sexual difficulties and were anchored at "absent" and "severe".

### 5.2.5. Experimental Interventions

#### *5.2.5.1. Antidepressant medication*

Following the completion of the initial assessment session all participants were provided with seven days supply of either citalopram 20 mg or lactose placebo capsules. The capsules were identical in appearance and participants were instructed to take them every morning, starting on the following morning and including the day of the post-treatment assessment.

#### *5.2.5.2. Attentional bias training*

The attentional bias training task used was a face based visual-probe training, similar to that described in Chapter Two. The decision to use faces rather than words during training was based on the recent publication of studies in which repeated face based training was found to be effective (Amir, Beard et al. 2009; Amir, Beard et al. 2009). The training task employed a total of 96 trials which were split into two blocks of 48 trials each. The blocks differed in the duration of time the stimuli were presented (500 ms or 1000 ms). The facial stimuli used in

training were taken from a range of sources (Matsumoto and Ekman 1988; Lundqvist, Flykt et al. 1998; Tottenham, Tanaka et al. 2009) and were split into two sets of images; for a given participant, one set was used in the initial assessment and the following training sessions with the second set being used for the final assessment session (sets were allocated at random). The stimuli displayed positive (happy), neutral or negative (fearful and angry) expressions. This resulted in three possible face pair types; positive-neutral, positive-negative, negative-neutral. Equal numbers of these pair-types were presented in each block. Images subtended a vertical visual angle of ≈11° and were separated by ≈3.5°. Following the initial assessment session, participants completed a single training session, under the supervision of an experimenter to ensure that the requirements of the task were understood. All participants then went on to complete the training sessions twice daily at home using a laptop computer which was supplied to them for the duration of the study. Administration of cognitive bias training at home as opposed to in the laboratory was selected as twice daily visits to the laboratory were felt likely to be unacceptable to participants and home based training has been demonstrated to be effective in a number of previous studies (See, Macleod et al. 2009; Blackwell and Holmes 2010).

In the positive training condition the probes were presented in the location of the relatively more positive face on 87.5% of trials (the probe was presented at the alternative location on the remaining 12.5% of trials allowing assessment of attentional bias during the training task itself). In the control training task, the location of the probe was random, that is there was no relationship between the expression of the face and the location of the probe (the control task therefore should not encourage either a positive or negative bias).

*5.2.5.3. Participant compliance*

Participant compliance with both the interventions (i.e. training and medication) was encouraged during the initial assessment session by explicitly planning the times at which training would be completed and medication taken. Further, the experimenters contacted all participants on two occasions during the treatment week to ensure that training sessions had been completed and medication taken. It was stressed to participants that completion of training would be recorded by the laptop and that a blood test would be taken following treatment to measure drug levels.

## 5.2.6. Assessment tasks

### *5.2.6.1. Visual-probe tasks; assessing attentional bias (MacLeod, Mathews et al. 1986)*

Word and face based visual-probe tasks were used to assess attentional bias before and after treatment. The tasks were similar to the training task in that attention to positive, negative and neutral stimuli were assessed. The word stimuli were taken from a range of previously published sources (Mathews, Mogg et al. 1989; Dozois and Dobson 2001; Dudley, O'Brien et al. 2002; Taghavi, Dalgleish et al. 2003; Lim and Kim 2005; Dannlowski, Kersting et al. 2006; Grant and Beck 2006) and, as described above, were split into two sets with one set being used for each testing session. In both tasks, the probe location was balanced between negative and positive stimuli and deployment of attention was assessed using blocks of trials in which the stimuli were presented for either 100 ms, or for the durations used in training (500 and 1000ms). This resulted in a total of 144 trials per task.

### *5.2.6.2. Emotional word categorisation task; assessing emotional bias for self referential information (Harmer, Shelley et al. 2004)*

Sixty personality characteristics selected to be extremely disagreeable (e.g. domineering, untidy, hostile) or agreeable (cheerful, honest, optimistic) (taken from Anderson 1968) were presented on the computer screen for 500 ms. These words were matched in terms of word

length, ratings of frequency and meaningfulness (Harmer, Shelley et al. 2004). Volunteers were asked to categorise these personality traits as likable or dislikeable as quickly and as accurately as possible. Specifically, they were asked to imagine whether they would be pleased or upset if they overheard someone else referring to them as possessing this characteristic, so that the judgment was in part self-referring. Reaction times for correct identifications were computed for this task.

*5.2.6.3. Emotional memory task; assessing memory bias (Harmer, Shelley et al. 2004)*

Participants were asked to recall as many of the personality trait words from the emotional word categorisation task above as possible (Harmer, et al., 2004). This task, therefore, allows the assessment of incidental memory for positive and negative characteristics (number of items correctly recalled for the likeable and dislikeable word conditions). Recognition memory was then assessed by asking volunteers to indicate whether they recognised the word for each item on a list containing the 60 targets along with 60 matched distractors (30 likeable, 30 dislikeable), providing an alternative assay of emotional memory. The relevant outcome from these tasks is the number of positive and negative words correctly recalled or recognised.

*5.2.6.4. Facial expression categorisation task; assessing interpretative bias (Harmer, Shelley et al. 2004)*

The facial expression recognition task features six basic emotions (happiness, surprise, sadness, fear, anger and disgust) taken from the Pictures of Affect Series (Ekman and Friesen 1976), which have been morphed between each prototype and neutral (Young, Rowland et al. 1997). These stimuli had not been used in the training task. Four examples of each emotion, at each intensity, were presented (total of 10 characters). Each face was also given in a neutral expression, giving a total of 250 stimuli presentations. The facial stimuli were

presented on a computer screen (in a random order) for 500 ms and replaced by a blank screen. Participants were required to identify the emotional expression on the face by pressing a labelled key on the keyboard. The outcome for this task is the accuracy when identifying the various emotions. Each participant was asked to respond as quickly and as accurately as possible.

*5.2.6.5. Emotional challenge tasks; assessing emotional reactivity to external events (Clark 1983; Moberly and Watkins 2006)*

In order to assess emotional reactivity to external events participants were required to complete both a mood induction task (Velten 1968; Clark 1983) and, separately, a social failure task (Moberly and Watkins 2006). During the mood induction task (Holmes, Lang et al. 2009), participants were required to listen to sad music (*Russia under the Mongolian Yoke* by Prokofiev played at 50% normal speed) while reading depressive statements on the computer screen ("It seems such an effort to do anything"). Such procedures have been reliably shown to produce increased depressive symptoms in non-clinical groups (Clark 1983) and to be sensitive to the effects of modification of interpretative bias (Holmes, Lang et al. 2009). During the failure task (Moberly and Watkins 2006) participants are required to identify a fourth word which is semantically linked to three other words which are presented on the screen. They are informed that the task measures "creative intelligence" and, following the task, their responses are marked and their overall score is fed back to them by the experimenter. The task is extremely difficult resulting in participants performing more poorly than expected and causing a reliable increase in negative symptoms (McFarlin and Blascovich 1984; Moberly and Watkins 2006). All participants completed the mood induction followed by the failure task. State measures of negative mood (negative-PANAS)

and anxiety (state-STAI) were taken before and after each task with the critical outcome being the change in these scores induced by the tasks.

### 5.2.7. Statistical Analysis

In order to test the competing hypotheses regarding the interaction of interventions on task outcomes, mixed model ANOVAs were constructed with the between subject factors of medication (citalopram, placebo) and training (positive, neutral). This resulted in estimates of the effect of each active interventions (i.e. main effects of training or medication) and, critically, an interaction term (medication x training) which assess for the presence of positive or negative interactions. The within subject factors differed between the tasks analysed. For the visual-probe tasks these were time (before treatment, after treatment), stimulus pair-type (positive-neutral, positive-negative, negative-neutral) and probe duration (100 ms, 500 ms, 1000 ms). The face categorisation task had a single within subject factor, face valence (fear, anger, surprise, happy, disgust, sad). Both the emotional word categorisation and memory tasks required a single within subject factor, word valence (positive, negative). Baseline questionnaire and demographic measures were analysed using univariate ANOVAs. Changes in self report scales (including both the changes occurring across the week of treatment and changes in momentary mood measures induced by the emotional challenge tasks) were analysed using univariate ANCOVAs in which the difference scores of the appropriate scales were entered as the dependent variable with the preceding score for the same variable entered as an additional covariate to account for any baseline difference between groups. Separate analyses of the 7 possible side effects were conducted, with the results being corrected using the bonferroni method for multiple comparisons.

Vigilance scores were calculated for the visual-probe tasks (see Chapter Two) such that a more positive vigilance score indicates increased attention towards the relatively more positive stimulus in any given pair (e.g. for the faces visual-probe this would be the happy face in the positive-neutral and positive-negative trials or the neutral face during negative-neutral trials). For all reaction time data median reaction time for correct trials was used in order to reduce the influence of outliers (MacLeod, Rutherford et al. 2002 and Chapter Two). Interpretation of the visual probe data, which had been collected under laboratory conditions during the assessment sessions, was complemented by plotting the vigilance scores from each training session (see section 5.2.5.2) in order to produce a graphical representation of the change of attentional function across training.

## 5.3. Results

### 5.3.1. Demographic, compliance, side effect and baseline data

The groups were well matched in terms of age, gender and baseline scores on mood and anxiety measures (see Table 5.1) indicating that randomisation had been successful. Compliance with the training regime was high with no participant missing more than 3 (out of 15) sessions and no difference between groups on the number of sessions completed [all $p > 0.13$]. There were no significant effects of group on changes in any of the side effects as measured by VAS [all corrected $p > 0.6$].

**Table 5.1: Summary of demographic and baseline data from all four groups**

| | Placebo with neutral training (n = 16) | Placebo with positive training (n = 15) | Citalopram with neutral training (n = 15) | Citalopram with positive training ( n = 16) | $p$[a] |
|---|---|---|---|---|---|
| **Age Mean (SD)** | 21.6 (2.4) | 21.1 (1.8) | 20.9 (2.6) | 21.1 (2.4) | 0.52 |
| **Gender F:M** | 8:8 | 8:7 | 8:7 | 8:8 | 0.80[b] |
| **BDI Mean (SD)** | 3.7 (2.9) | 3.9 (5.2) | 2.6 (2.7) | 2.9 (2.4) | 0.24 |
| **Trait-STAI Mean (SD)** | 39 (7.6) | 35 (8) | 34.3 (6) | 35.7 (7.7) | 0.19 |
| **Neg-PANAS Mean (SD)** | 13 (3.2) | 12.8 (6) | 11.9 (1.6) | 11.7 (2.3) | 0.23 |
| **Pos-PANAS Mean (SD)** | 32.5 (7.3) | 33.1 (5.9) | 34.4 (6.9) | 34.1 (6.6) | 0.40 |
| **State-STAI Mean (SD)** | 29.6 (6.9) | 29.8 (10.1) | 29.2 (7) | 29.4 (6.1) | 0.80 |

[a] The *p* value reported is the lowest from either of the main effects (medication, training) or the interaction term (medication x training)
[b] Analysis performed using logisitic regression model. All other analysis performed using univariate ANOVA.
BDI=Beck Depression Inventory, STAI=Spielberger State-Trait Anxiety Inventory, PANAS=Positive and Negative Affect Scale

## 5.3.2. Did the interventions alter attentional function?

### *5.3.2.1. Attention to faces*

As expected, positive attentional bias training, when compared to neutral training, lead to a general increase in attentional bias towards positive faces [see Figure 5.2a and Table 5.2. Main effect of training condition; $F(1,58) = 4.6$, $p = 0.04$]. This effect was not modified by the duration at which the stimuli were presented (100 ms, 500 ms, 1000 ms), the face pair types which were presented (positive-neutral, positive-negative, negative-neutral) or any

interaction of these factors [all $p > 0.3$]. These results indicate that the training induced a generalised positive bias to novel faces, which was also evident specifically within the 100 ms trials which were not included in training [$F(1,58) = 4.0$, $p = 0.05$]. In contrast to the effect of attentional training, no effect of medication group was found on attentional bias to faces [main effect of medication; $F(1,58) = 0.36$, $p = 0.55$] and there was no evidence of an interaction between the medication and attentional training [$F(1,58) = 0.6$, $p = 0.44$]. Inspection of the vigilance scores obtained during the training task itself (see section 5.2.5.2 for a description of the task) reveals a similar effect of the interventions, with a clear training effect emerging over time only for the groups who undertook positive training (Figure 5.2b).

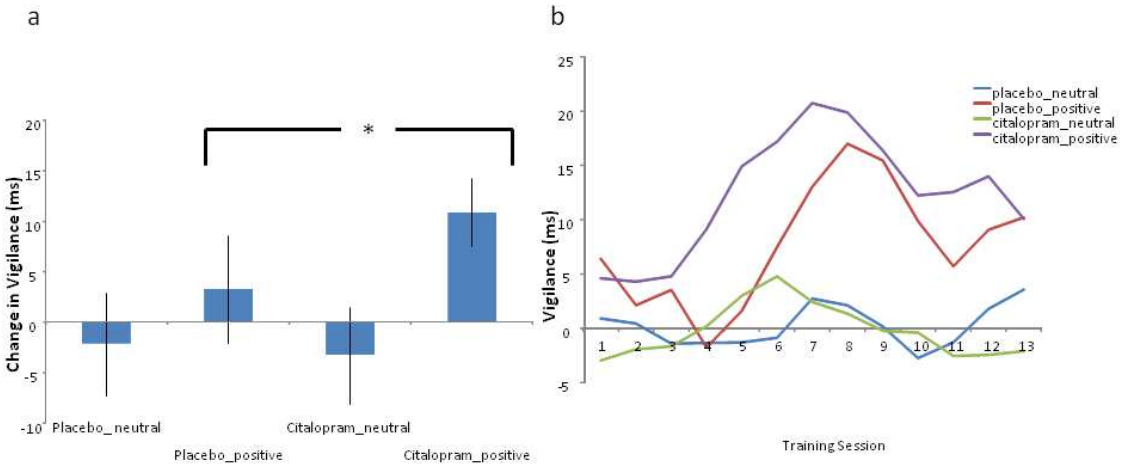



**Figure 5.2: Effect of the interventions on attention to faces. a) change in vigilance as measured using the faces visual-probe task in the two assessment sessions. Analysis (see section 5.3.2.1) reveals a main effect of training group only. Bars represent mean (SEM) change in the vigilance scores. * = $p < 0.05$. b) Change in vigilance as measured during the training task (see section 5.2.5.2). Each training session is represented on the x axis. Data have been smoothed using a sliding average of two sessions. Vigilance is calculated such that a more positive number indicates vigilance towards the positive faces (see section 5.2.7).**

*5.3.2.2. Attention to words*

There was no evidence for the generalisation of the training effect from faces to words (Table 5.2). Specifically the change in attentional bias towards word stimuli was not altered by bias training [$F(1,58) = 0.45$, $p = 0.5$], medication [$F(1,58) = 1.6$, $p = 0.21$] or their interaction [$F(1,58) = 0.2$, $p = 0.7$].

## 5.3.3. Did the interventions alter other measures of cognitive bias?

### *5.3.3.1. Emotional memory*

Following the intervention phase, a positive memory bias was apparent across all participants in the emotional recognition task with a greater number of positive words being recognised than negative words [$F(1,58) = 165$, $p < 0.001$]. A training x medication interference effect, across both positive and negative words, was also observed for the accuracy data [$F(1,58) = 10.5$, $p = 0.002$]. Critically however, this interaction was modified by the emotion of the word [training x medication x emotion; $F(1,58) = 5.1$, $p = 0.03$] such that each intervention administered singly produced a relatively positive bias in recognition whereas the combination of interventions did not (see Figure 5.3a and Table 5.3). This emotion specific memory effect was driven by a pronounced training x medication interaction for negative words [$F(1,58) = 12.6$, $p = 0.001$], with a similar pattern being found to a significantly reduced extent for the recognition of positive words [$F(1,58) = 4.4$, $p = 0.04$]. As can be seen from Figure 5.3b each intervention on its own produced a significant reduction in the recognition of negative words [placebo_neutral vs. placebo_positive; $t(29) = 2.8$, $p = 0.01$. placebo_neutral vs. citalopram_neutral $t(29) = 3.5$, $p = 0.001$], whereas the combination of interventions produced no change from baseline [placebo_neutral vs. citalopram_positive; $t(30) = 1$, $p = 0.3$]. In fact, the combined intervention group displayed increased recognition memory for negative words when compared to the individual intervention groups [citalopram_positive vs. placebo_positive; $t(29) = 1.8$, $p = 0.09$. citalopram_positive vs. citalopram_neutral $t(29) = 2.3$, $p = 0.03$]. In short, while the individual interventions of medication or training produced a positive bias in recognition memory, the combination of interventions produced an interference effect.

Analysis of the emotional recall task (Table 5.3) confirmed the general positive bias found across all participants [$F(1,58) = 11.9$, $p = 0.001$] but found no effect of either treatment [main effect of training; $F(1,58) = 0.77$, $p = 0.38$. main effect of medication; $F(1,58) = 0.03$, $p = 0.86$] or their interaction [$F(1,58) = 0.14$, $p = 0.76$].

*5.3.3.2. Emotional word categorisation*

The relative speed with which participants classified positive and negative self referential words (Table 5.3) was influenced by a borderline significant interaction between the interventions [emotion of word x medication x training; $F(1,58) = 3.8$, $p = 0.056$]. As can be seen from Figure 5.3c all groups responded to the positive word more quickly than the negative word [$F(1,58) = 28.2$, $p < 0.001$]. When considering the individual differences between the placebo_neutral group and the singly administered interventions, the results did not reach statistical significance, although the effect of citalopram did reach trend level [placebo_neutral vs. placebo_positive; $F(1,29) = 0.74$, $p = 0.35$. placebo_neutral vs. citalopram_neutral; $F(1,29) = 4.0$, $p = 0.055$]. There was no evidence that the initial interaction was driven specifically by either word valence as, when considering each valence separately, the medication x training interaction was non-significant [positive words; $F(1,58) = 0.28$, $p = 0.6$. negative words; $F(1,58) = 0.15$, $p = 0.7$]. Thus the emotional categorisation task revealed a trend level interference effect which appeared similar in structure to that found in the recognition task.

**Table 5.2: Raw group reaction times for trials from the visual-probe tasks. In order to limit the number of results displayed the reaction times have been averaged over the dimensions of the tasks which did not influence the results (i.e. the duration of stimuli presentation and the pair type presented). The results presented are the mean (standard deviation) of the median reaction times for each participant.**

| Time | Before Intervention | | | | After Intervention | | | |
|---|---|---|---|---|---|---|---|---|
| Group | Placebo_ neutral | Placebo_ positive | Citalopram_ neutral | Citalopram_ positive | Placebo_ neutral | Placebo_ positive | Citalopram_ neutral | Citalopram_ positive |
| Positive Words | 497 (60) | 519 (54) | 504 (54) | 496 (43) | 442 (54) | 474 (60) | 439 (32) | 452 (55) |
| Negative Words | 494 (54) | 518 (51) | 499 (50) | 496 (43) | 438 (48) | 471 (43) | 444 (32) | 455 (52) |
| Positive Faces | 464 (49) | 493 (45) | 475 (52) | 474 (54) | 410 (43) | 434 (35) | 408 (33) | 426 (58) |
| Negative Faces | 469 (51) | 489 (47) | 477 (50) | 472 (58) | 412 (44) | 434 (34) | 407 (35) | 434 (61) |

**Table 5.3: Raw group data from the categorisation, recognition and recall tasks. For the categorisation tasks the mean (standard deviation) of the median reaction times are displayed. For the other two tasks the mean (standard deviation) of the number of correctly recognised or recalled words are displayed.**

| Group | Placebo_ neutral | | Placebo_ positive | | Citalopram_ neutral | | Citalopram_ positive | |
|---|---|---|---|---|---|---|---|---|
| Word Valence | Positive | Negative | Positive | Negative | Positive | Negative | Positive | Negative |
| Categorisation [mean (SD) | 774 (147) | 846 (142) | 793 (128) | 844 (125) | 806 (211) | 817 (185) | 782 (140) | 844 (123) |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **reaction time]** | | | | | | | | |
| **Recognition [mean (SD) recognised]** | 26.2 (2.1) | 22.5 (2.4) | 24.8 (3.5) | 19.1 (4.2) | 23.8 (2.2) | 18.9 (3.2) | 25.2 (2.6) | 21.5 (3.2) |
| **Recall [mean (SD) recalled]** | 4 (3.2) | 1.1 (1.3) | 3.4 (1.7) | 1.3 (1.3) | 4.2 (1.7) | 1.3 (1.1) | 3.9 (2.6) | 0.8 (0.9) |

**Table 5.4: Data from the facial expression categorisation task. Results display the mean (standard deviation) of the proportion of faces correctly identified.**

| Group | Placebo_ neutral | Placebo_ Positive | Citalopram_ neutral | Citalopram_ positive |
|---|---|---|---|---|
| **Anger** | 0.54 (0.11) | 0.49 (0.09) | 0.50 (0.09) | 0.47 (0.12) |
| **Disgust** | 0.53 (0.17) | 0.44 (0.18) | 0.50 (0.15) | 0.45 (0.15) |
| **Fear** | 0.46 (0.15) | 0.46 (0.14) | 0.42 (0.07) | 0.43 (0.12) |
| **Happy** | 0.65 (0.09) | 0.68 (0.10) | 0.61 (0.11) | 0.63 (0.12) |
| **Sad** | 0.49 (0.12) | 0.46 (0.10) | 0.50 (0.15) | 0.42 (0.14) |
| **Surprise** | 0.56 (0.12) | 0.46 (0.08) | 0.50 (0.13) | 0.50 (0.12) |
| **Neutral** | 0.66 (0.17) | 0.75 (0.22) | 0.71 (0.18) | 0.70 (0.18) |

**Table 5.5: Raw group data from the emotional challenge tasks. Numbers indicate the mean (standard deviation) of the relevant scale.**

| Group | Placebo_ neutral | | | Placebo_ positive | | | Citalopram_ neutral | | | Citalopram_ positive | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Time point** | **Before mood induction** | **Between mood induction and social failure task** | **After social failure task** | **Before mood induction** | **Between mood induction and social failure task** | **After social failure task** | **Before mood induction** | **Between mood induction and social failure task** | **After social failure task** | **Before mood induction** | **Between mood induction and social failure task** | **After social failure task** |
| **Negative-PANAS** | 12.6 (4.6) | 14.6 (6.7) | 14.6 (6.1) | 12 (3.8) | 14.1 (6.0) | 12.9 (4.7) | 11.5 (1.6) | 10.9 (1.2) | 12.8 (3.4) | 11.7 (1.6) | 12.4 (3.0) | 12.4 (3.4) |
| **State-STAI** | 32.9 (9.9) | 36.4 (11.6) | 36.9 (11.8) | 30.3 (10.6) | 34.5 (13.0) | 34.7 (12.6) | 30 (6.9) | 31.5 (7.6) | 33.1 (8.8) | 31.1 (4.8) | 33.2 (7.9) | 34.8 (8.3) |

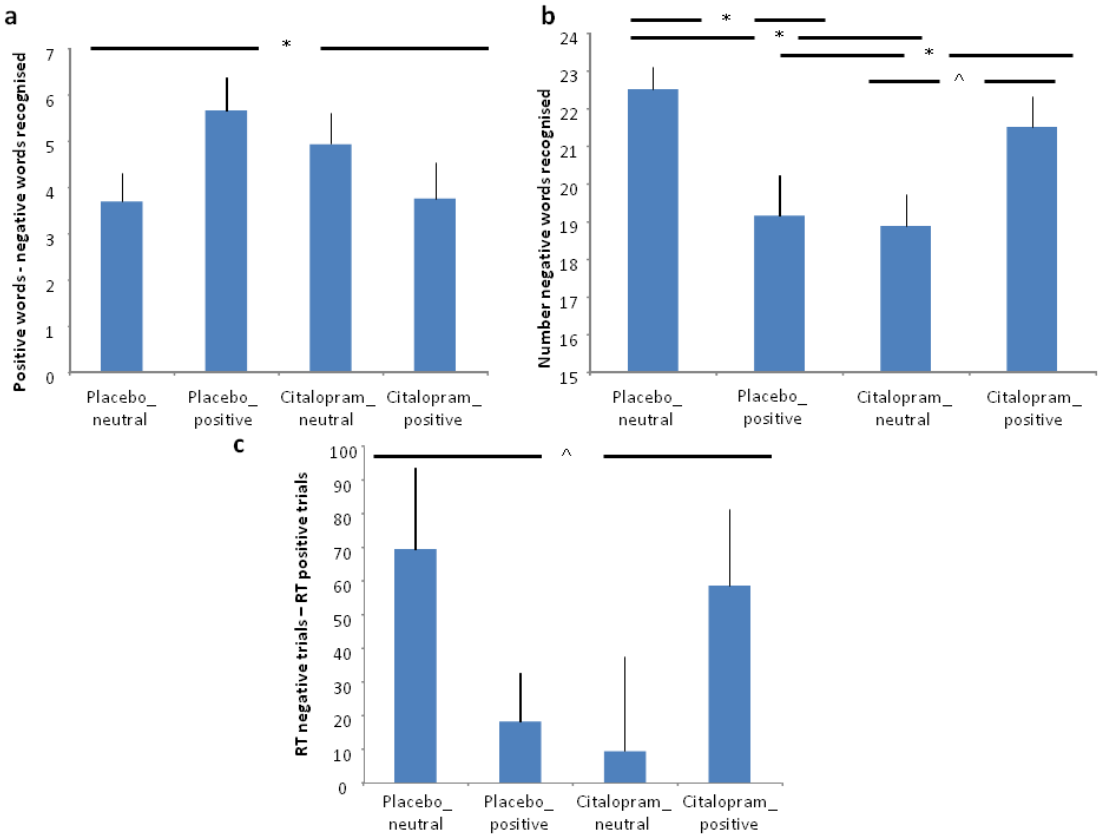



**Figure 5.3: Effects of experimental interventions on measures of cognitive bias. a) relative accuracy when remembering positive vs. negative words. An interference pattern is seen; both interventions, administered singly, produce a relative increase in positive memory bias, whereas the combination treatment does not. b) raw scores for number of negative words recognised. Both citalopram and positive training significantly decrease recognition of negative words whereas the combination of interventions has no effect. c) reaction time to classify negative vs. positive self-referential words. A trend towards a negative interference pattern is seen, with each intervention administered singly reducing the relative reaction time difference between negative and positive words whereas the combination of interventions again has very little effect. All graphs display mean (SEM). * = p < 0.05, ^ = p < 0.1.**

### *5.3.3.3. Emotional face categorisation*

There were no emotion specific effects of either treatment [training; $F(5,290) = 1.7$, $p = 0.14$. Medication; $F(5,290) = 0.26$, $p = 0.94$] or their interaction [$F(5,290) = 0.8$, $p = 0.55$] on accuracy when categorising the emotional faces (Table 5.4).

### *5.3.3.4. Are the categorisation and memory effects related?*

The similarity between the interference effects found in the emotional recognition and words categorisation tasks suggests that these effects may be linked. Indeed, the word categorisation task may be considered the encoding phase of the memory task, as the latter

tests whether participants can recognise the words presented in the former. Consistent with this hypothesis, a significant correlation was found between the reaction time difference scores on the categorisation task (NB RT for negative words – RT for positive words) and memory for the negative words [$r(62) = 0.31$, $p = 0.01$] and more specifically between reaction time for the negative words and subsequent memory for the same words [$r(62) = 0.32$, $p = 0.01$, see Figure 5.4]. Thus the longer the participant took to categorise the negative words, the more likely they were to recognise them.

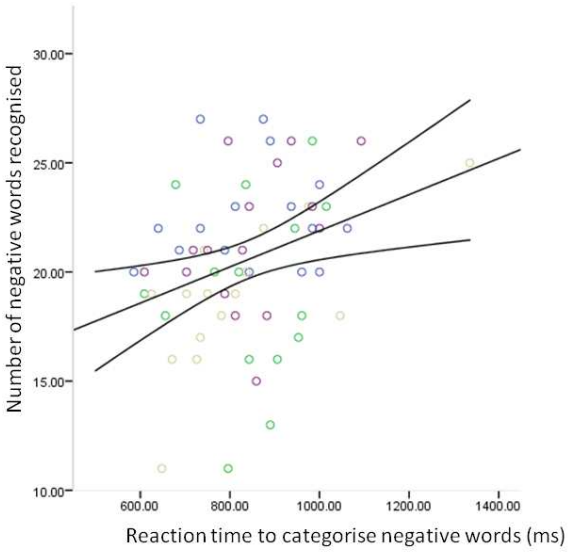



**Figure 5.4: Scatter plot showing correlation between reaction time during word categorisation and subsequent memory for negative words, across all participants. A fitted regression line with 95% confidence intervals has been superimposed. Correlation remains significant after removing outlying data (recognition <15 and reaction time > 1200). Blue circles = placebo_neutral, green circles = placebo_positive, grey circles = citalopram_neutral, maroon circles = citalopram_positive.**

## 5.3.4. Did the interventions alter baseline measures of mood or anxiety?

Consistent with the assumption that the interventions do not act to directly alter mood or anxiety, there was no effect of either intervention or their interaction on change in any of the baseline measures listed in Table 5. [all $p > 0.14$].

## 5.3.5. Did the interventions alter emotional reactivity to external stressors?

### *5.3.5.1. Mood induction task*

The increase in depressive mood caused by the mood induction procedure was assessed using the negative-PANAS scale which was administered both before and after participants completed the mood induction (Table 5.5). This revealed a protective effect of citalopram, with participants who had taken citalopram experiencing a significantly smaller increase in negative mood than those who had received placebo [main effect of medication; $F(1,57) = 6.0$, $p = 0.02$]. In contrast there was no effect of attentional bias training [main effect of training; $F(1,57) = 1.0$, $p = 0.31$] and no interaction between the interventions [$F(1,57) = 0.58$, $p = 0.45$]. Post hoc analyses of the significant medication effect revealed that there were no differences in negative mood between the groups receiving placebo and those receiving citalopram before the mood induction [(placebo_neutral + placebo_positive) vs. (citalopram_neutral + citalopram_positive); $t(60) = 0.89$, $p = 0.38$] but those receiving citalopram reported lower levels of negative affect following the procedure [(placebo_neutral + placebo_positive) vs. (citalopram_neutral + citalopram_positive); $t(60) = 2.2$, $p = 0.03$]. Similarly, across the procedure, participants receiving placebo reported a significant increase in negative symptoms [paired t-test, (placebo_neutral + placebo_positive); $t(30) = 3.2$, $p = 0.004$], whereas those who received citalopram did not [paired t-test, (citalopram_neutral + citalopram_positive); $t(30) = 0.26$, $p = 0.8$]. Analysis of anxiety symptoms using the state-STAI revealed a general increase in anxiety across all participants [$F(1,57) = 7$, $p = 0.01$], with no effect of either treatment [main effect of training; $F(1,57) = 0.5$, $p = 0.80$. main effect of medication; $F(1,57) = 1.2$, $p = 0.29$] or their interaction [$F(1,57) = 0.04$, $p = 0.85$].

*5.3.5.2. Social failure task*

The failure task caused an increase in negative affect (Table 5.5) as assessed using the negative-PANAS across all participants [$F(1,57) = 11.4$, $p = 0.001$]. This was modified by a trend level tendency for positive attentional training to protect against the increase of

negative mood [main effect of training; $F(1,57) = 3.9$, $p = 0.054$] with no effect of medication [main effect of medication; $F(1,57) = 1.15$, $p = 0.29$] or of the treatment interaction [$F(1,57) = 0.02$, $p = 0.89$]. Additionally the baseline measure of mood which, as described above, differed significantly between the medication groups, showed an inverse relationship to mood change [$F(1,57) = 11.4$, $p = 0.001$] such that the higher the level of pre-task negative symptoms, the smaller the increase in those symptoms. Continuing to control for this baseline difference, there was a significant increase in negative symptoms across the task for those participants who received neutral training [change in negative-PANAS in (placebo_neutral + citalopram_neutral) controlling for initial score; $F(1,28) = 5.8$, $p = 0.02$] whereas there was a small but significant reduction in negative affect for the positive training group [change in negative-PANAS in (placebo_positive + citalopram_positive) controlling for initial score; $F(1,28) = 6.1$, $p = 0.02$]. Similarly there was a trend towards the positive training group reporting lower negative affect after the task [(placebo_neutral + citalopram_neutral) vs. (placebo_positive + citalopram_positive) controlling for initial score; $F(1,57) = 3.9$, $p = 0.054$] (NB a controlled assessment of the pre-task scores is not possible as the same scores are used as the covariate).

The failure task induced a trend towards an increase in anxiety symptoms, as measured by the state-STAI, across all participants [$F(1,57) = 3.0$, $p = 0.09$] with no effect of either treatment [main effect of training; $F(1,57) = 0.01$, $p = 0.93$. main effect of medication; $F(1,57) = 0.58$, $p = 0.45$] or their interaction [$F(1,57) = 0.05$, $p = 0.82$] on this.

*5.3.5.3. Are the protective effects of the interventions specific to a particular form of emotional challenge?*

The analysis reported above indicates that citalopram was protective during the mood induction procedure but not the social failure task whereas positive bias training tended to have the opposite pattern of efficacy. In order to test whether there was statistical evidence for this specificity a repeated measures ANOVA was performed on the changes in the negative-PANAS scores for both tasks. A single within subject factor of type of emotional challenge was included (mood induction task, social failure task) as were the between subject factors of medication and training groups (NB it was not possible to control this analysis for both baseline measures as the change score over the mood induction task is equivalent to the difference between these baseline measures). This analysis revealed a significant task x medication interaction [$F(1,58) = 7.6$, $p = 0.008$] and a trend towards a task x training interaction [$F(1,58) = 3.3$, $p = 0.075$] indicating that the effects of medication were specific to the mood induction procedure whereas training tended to specifically influence mood across the social failure task.

### 5.3.6. Are the cognitive effects of the interventions related to their effects on emotional reactivity?

Cognitive models of treatment effects in the emotional disorders (Harmer 2008) suggest that altering the processing of emotional information should causally impact on the emotional experience of patients. This would predict that, within an active treatment group, the cognitive effects of an intervention should correlate with its effect on emotional reactivity. Consistent with this a significant positive correlation within the citalopram treated participants was found between the number of negative words recognised during the memory task and the subsequent increase in the negative-PANAS score induced by the mood induction [$r(31) = 0.44$, $p = 0.01$]. There was no correlation between memory task scores and

response to the social failure task [$r(31) = -0.01$, $p = 0.94$] and the positive attentional bias training group showed no correlation for either of the two challenge tasks [mood induction; $r(31) = -0.29$, $p = 0.88$. social failure task; $r(31) = -0.01$, $p = 0.94$].

## 5.4. Discussion

The current study is the first to explore the interaction between attentional bias training and an antidepressant medication, citalopram, on both cognitive function and emotional reactivity to external challenge. Concurrent administration of the active interventions was found to produce an interference effect on a measure of emotional memory with a similar effect being found at a trend level during the encoding of the same words. However, assessment of emotional reactivity to external challenge did not reveal interference effects, in this case there was evidence for selective effects of the interventions in response to different forms of challenge; whereas citalopram protected participants against the negative mood induced by the mood induction procedure, positive attentional training tended to protect against worsening mood induced by a social failure task.

### 5.4.1. The effects of the interventions on emotional memory and interpretation

The finding that the citalopram_neutral group recognised fewer negative words than the placebo_neutral group is consistent with the previously reported positive effect of citalopram on emotional memory (Harmer, Shelley et al. 2004). The current study demonstrates that this cognitive effect is not specific to antidepressant medication however, as positive attentional bias training was found to have a similar effect in the same task. The importance of these findings is that they suggest that the emotional memory task may be used to assess the combination of drug and cognitive interventions within a single participant. Given this, the reduced memory effect demonstrated in the combined intervention group suggests that,

although the gross effects of each intervention in clinical populations appears similar (see Chapter One), they combine destructively at a cognitive level. In the introduction, the generally disappointing results from clinical meta-analyses of combination antidepressant and psychotherapy treatments (Foa, Franklin et al. 2002; Pampallona, Bollini et al. 2004; Cuijpers, van Straten et al. 2009) was reviewed. If it is assumed that attentional bias training mimics one facet of more complex psychotherapies (see Chapter One) then the results from the memory task raises the possibility that the poor performance of combination treatments may, at least in part, result from interference between cognitive and pharmacological interventions at the cognitive level.

The trend towards a similar pattern of interference in the categorisation task provides evidence on the component processes of the memory effect described above. Specifically, this finding suggests that the memory effect may be due to the influence of the interventions during encoding, a proposal supported by the positive correlation between reaction time during categorisation and the subsequent number of negative words recognised. A tempting interpretation of these results is that the time spent encoding provides a Stroop-like index of the attentional allocation to the stimuli; in other words the interventions influenced the attentional resources participants deployed to the positive vs. negative words, with increased allocation of attention allocated to the word being associated with an increased reaction time when categorising the word. If true, this would suggest that the observed memory effect arises as a direct consequence of the attentional effects of the interventions. However, this conclusion would seem to be at odds with the results from the visual-probe tasks, which directly assess the deployment of attention. In these tasks attentional bias training was found to influence attention to novel faces, whereas medication did not influence this measure and

neither treatment altered attentional bias to words. The lack of a medication effect on these measures was surprising given that a number of previous studies have demonstrated a relatively positive attentional effect of citalopram (Browning, Reid et al. 2007; Murphy, Yiend et al. 2009). The most obvious reason for the failure to replicate these results is that the training procedure completed by all participants in the current study involved repeated administration of a version of the visual-probe task and thus, by the time the effects of medication were assessed, repeated practice of the visual-probe procedure may have reduced the sensitivity of the tasks to detect the influence of the medication. This difficulty may be circumvented in future work by including an alternative assessment of attention (e.g. using a Stroop type task; Williams, Mathews et al. 1996) which participants had not practised.

The lack of effect of either intervention on the words visual-probe task is similar to the finding from Chapter Two in which a single episode of face training did not alter attention to words. It is difficult to argue that the training intervention in the current study was not effective or did not generalise, as it influenced performance on a measure of a distinct cognitive system, emotional memory (for word stimuli). Thus it may be that the difficulty in demonstrating attentional effects with word stimuli was due to the words themselves being less emotionally potent than pictures of faces (a similar suggestion is made by Mogg and Bradley 1998) and therefore less able to influence reaction time during completion of the visual-probe task.

Interestingly, while generalisation from faces to words was not demonstrated, there was evidence of generalisation of the training effect across stimulus duration (c.f. Chapter Two) as the training task itself presented stimuli for 500 and 1000 ms whereas the influence of training was found across both these durations and when stimuli were presented for 100 ms.

This is consistent with the proposal in Chapter Two that repeated training would lead to a more generalised attentional effect.

In summary, performance on the emotional memory task was sensitive to the individual effects of both antidepressant medication and attentional bias training. Combining the interventions was found to erode their individual effects. The influence of the interventions on memory function may have arisen consequent to their attentional effects, a proposal which may be tested in future studies using convergent measures of attentional deployment.

### 5.4.2. The effects of the interventions on emotional reactivity to external challenge

In contrast to the interference between the interventions found in the emotional memory task, the results from the emotional challenge tasks suggested that the interventions displayed distinct protective effects when participants were faced with challenging situations. Interestingly the effects of the interventions were not simply independent; in other words, there was no evidence of non-interacting treatment effects on a *single* outcome measure. Rather, the different interventions appeared to be specifically beneficial in response to particular forms of external challenge. Citalopram reduced the negative response to the mood induction procedure but had no effect on the social failure task whereas positive attentional bias training tended to be protective across the social failure task without influencing the response to mood induction.

The finding that positive attentional bias training tended to protect against the negative affect induced by social failure is consistent with previous studies which have demonstrated a similar effect using a range of training tasks (MacLeod, Rutherford et al. 2002; Watkins, Moberly et al. 2008; Amir, Beard et al. 2009; Krebs, Hirsch et al. 2010). The observation that

citalopram protects against the effect of negative mood induction is novel. While previous studies do suggest that serotonergic manipulation may influence mood; for example, tryptophan depletion has been reported to increase negative mood in previously depressed patients (Smith, Fairburn et al. 1997), there has been no previous demonstration of a treatment-like protective effect of antidepressant medication on evoked mood.

It is not obvious why the protection afforded by the different interventions appeared to be specific to particular forms of external challenge. One possible explanation may be found in the very different nature of the two challenge tasks; the mood induction procedure encourages a generally negative affect by playing negative music and asking participants to simply read negative statements whereas the social failure task invokes a more complex response by setting up a tendency for participants to perform worse than they anticipate in a social situation. It may be, therefore, that citalopram protects against the type of non-specific negative experience evoked by the mood induction procedure, conceivably by altering the function of amygdala based circuitry (see Chapter One) in such situations whereas the positive attentional bias training is effective at reducing more nuanced reactions to complex situations, consistent with its influence on frontal function (Chapter Three). Regardless of the mechanisms underlying this specificity, the current results raise the intriguing clinical possibility that the treatment efficacy of pharmacological and cognitive interventions may be partly determined by the type of ongoing stressors experienced by patients.

### 5.4.3. Are the cognitive and emotional findings linked?

The rationale for the current study was based on cognitive models of the emotional disorders (Mathews and MacLeod 2005) and their treatments (Harmer 2008) which suggest that alteration in the processing of emotional information mediates the efficacy of the treatments.

In the context of the current study these models predict that the cognitive and emotional effects of the interventions should be correlated. The evidence in support of these predictions is somewhat mixed. Thus, while the finding of a positive correlation in the citalopram group, between the number of negative words recognised and the subsequent increase in negative affect induced by the mood induction, are consistent with the cognitive effects of the drug mediating its effect on mood, other results are more difficult to explain. In particular, a clear interference effect was found on the memory task with no evidence of interference on the emotional reactivity of participants. This discrepancy, when considered together with the effects of the interventions on specific external stressors discussed above, suggests that environmental factors may modify the expression of the cognitive biases induced by treatment interventions. Future modifications of the cognitive model of treatment mechanisms may therefore benefit from incorporating the effects of such treatment x environment interactions.

### 5.4.4. Limitations

A number of limitations to the current study must be considered. Firstly, it is not clear that mood induction effects in non-clinical populations can be straightforwardly extrapolated to clinical groups (Polivy and Doyle 1980). It would therefore be helpful to assess the effect of treatment interactions in further studies of clinical groups. Secondly, the current study has not replicated all the effects of citalopram reported in previous studies. In particular, the emotional face categorisation task has previously been found to be influenced by the medication (Harmer, Bhagwagar et al. 2003; Harmer, Shelley et al. 2004; Browning, Reid et al. 2007), although no effects were found in the current study. The reason for this

inconsistency is unclear, although it is possible that participants' response to the facial expressions were altered by the repeated exposure to facial stimuli during the training task.

### 5.4.5. Alternative treatment combinations

From a pragmatic perspective the current study has clearly been unable to demonstrate any evidence for what would have been the most promising clinical finding; a positive interaction between treatments. Given this it would be interesting to deploy the factorial design of the current study, which is particularly powerful when investigating treatment interactions, to other combinations of pharmaco and psychotherapies. For example, there is some evidence that D-Cycloserine is able to improve outcome in exposure based psychotherapies (Kushner, Kim et al. 2007; Storch, Merlo et al. 2007; Norberg, Krystal et al. 2008; Wilhelm, Buhlmann et al. 2008) suggesting that this pharmacotherapy may improve learning in other forms of psychotherapy such as attentional bias training. Alternatively, antidepressant medication may interact more constructively with alternative cognitive interventions, such as interpretative bias training (Mathews and MacLeod 2002; Blackwell and Holmes 2010) which encourages a more explicit bias alteration than that used in the current study.

### 5.4.6. Conclusion

Having investigated the mechanisms underlying attentional bias training in previous chapters, the current study explored the clinically relevant question of how this intervention interacts with a serotonergic antidepressant medication. The results indicate that the two forms of treatment interacted destructively on measures of cognitive function while producing distinct effects on emotional reactivity. In the following chapter a second clinically relevant question is addressed; do patients with bipolar disorder display attentional biases to emotional information and can attentional function in this clinical group be altered by training?

# Chapter 6

# Attentional Bias and its Modification in Patients with Bipolar Disorder

## 6.1. Introduction

In previous chapters the neural and behavioural effects of attentional bias training have been investigated in non-clinical populations. In the final data chapter of this thesis the behavioural effects of a single session of attentional bias training are assessed in a group of patients with a serious mental illness; bipolar disorder.

While the diagnosis of bipolar disorder requires episodes of raised mood (mania or hypomania) and is generally associated with episodes of depression (APA 1994), a range of other symptoms are commonly associated with the diagnosis. A striking example of this is the finding that anxiety disorders occur particularly frequently in patients with bipolar disorder, with the lifetime prevalence of any anxiety disorder having being estimated to be as high as 92% (Kessler 1999; McIntyre, Soczynska et al. 2006). Furthermore, the presence of a comorbid anxiety disorder is associated with a greater severity of bipolar illness, greater functional impairment and, in particular, with increased suicidality (Simon, Otto et al. 2004; McIntyre, Soczynska et al. 2006; Otto, Simon et al. 2006; Simon, Zalta et al. 2007). Importantly, the anxiety symptoms experienced by patients with bipolar disorder may often present outside the episodes of mania or depression (McIntyre, Soczynska et al. 2006) and thus, in theory, may be amenable to treatment when patients are "euthymic". Overall, these

results suggest that the reduction of anxiety may be a reasonable target for treatment in bipolar disorder. Such observations motivate the current study which provides an initial assessment of the utility of attentional bias training, which has previously been shown to reduce anxiety in other clinical groups (Amir, Beard et al. 2009; Amir, Beard et al. 2009; Schmidt, Richey et al. 2009).

Bipolar disorder has been repeatedly shown to be associated with neuropsychological deficits across a range of cognitive tests (Robinson, Thompson et al. 2006; Torres, Boudreau et al. 2007), with executive function, sustained attention and verbal memory appearing to be particularly affected (Clark and Sahakian 2006). In contrast, the evidence for varieties of emotional bias has been inconsistent. Considering specifically attention to emotional information, studies using the emotional Stroop task have tended to find either interference with the naming of depression relevant words, indicating attentional bias towards negative information (Bentall and Thompson 1990; French, Richards et al. 1996; Lyon, Startup et al. 1999; NB the first two studies recruited non-clinical, analogue populations), or simply a generalised slowing in response (Kerr, Scott et al. 2005). Studies utilising the visual-probe task have reported both a general avoidance of emotional words (Jongen, Smulders et al. 2007) and a bias towards threatening faces in children with dual diagnoses of bipolar disorder and an anxiety disorder (Brotman, Rich et al. 2007). Thus while there is some evidence for disturbance of attentional deployment to emotional stimuli in bipolar disorder the nature of the disturbance is not clear. One possible explanation for these inconsistent findings is provided by the study by Brotman and colleagues (2007) which reported a threat related attentional bias in bipolar children, but only when there was a comorbid anxiety disorder. These findings suggest that negative attentional bias in bipolar disorder may be associated

with the levels of anxiety experienced by patients rather than being a fundamental component of the illness itself. As the majority of previous studies have examined attentional bias towards "mania relevant" or "depression relevant" stimuli rather than the threatening stimuli which evoke attentional bias in anxious participants (see Chapter One), this anxiety-linked attentional bias may have been missed. The initial aim of the current study is therefore to assesses whether adult patients with bipolar disorder, when compared to healthy controls, display a similar attentional bias to that found in the anxiety disorders; that is, a bias towards briefly presented, threatening information, and whether attentional function is influenced by the levels of anxiety of the patients.

To date little work has examined treatment strategies specifically aimed at improving anxiety symptoms in bipolar disorder (McIntyre and Keck 2006; Kauer-Sant'Anna, Kapczinski et al. 2009). Treatment decisions are further complicated by concerns that antidepressant medication, a mainstay of the treatment of primary anxiety disorders (NICE 2007), may induce mania in bipolar patients (Keck, Strawn et al. 2006) and by the generally poor performance of psychotherapy in patients with the disorder (Colom, Vieta et al. 2003; Miklowitz, George et al. 2003; Lam, Hayward et al. 2005; Scott, Paykel et al. 2006; Lynch, Laws et al. 2010). Accordingly, the second aim of the current study was to provide an initial assessment of the effects of attentional bias training on a sample of patients with bipolar disorder. This was achieved by randomly assigning patients to either a single session of positive or neutral attentional bias training and then assessing the effects of training on both attentional function and, as in Chapter Five, emotional reactivity to external challenge.

Based on the epidemiological evidence linking anxiety to bipolar disorder it was predicted that patients with bipolar disorder would report greater levels of anxiety symptoms than non-

clinical controls. It was also predicted that attentional bias towards negative stimuli would be increased in the bipolar group when compared to the non-clinical controls, particularly those with a comorbid anxiety disorder. Lastly it was predicted that positive attentional bias training would decrease attentional bias to threat and would reduce the negative emotional reactivity to an external emotional challenge.

## 6.2. Methods

### 6.2.1. Overview of study

The study consists of two phases (see Figure 6.1); an initial case-control phase in which the baseline attentional biases of a group of patients with bipolar disorder was compared to a group of healthy controls which had been matched to the patient group in terms of average age, gender and years of education. This was followed by a second, randomised phase, in which the group of patients were randomly assigned to either a positive or neutral attentional bias training regime. Attentional bias to both words and faces presented for 100 ms and 500 ms was assessed at baseline. Following training attentional bias was reassessed to words and faces at 500 ms and emotional reactivity to an external emotional challenge was assessed. The order of testing (words vs. faces and 100 ms vs. 500 ms) was counterbalanced between participants. A booster session of training was completed between the two post training assessment tasks with a second booster session being completed before the emotional challenge task.

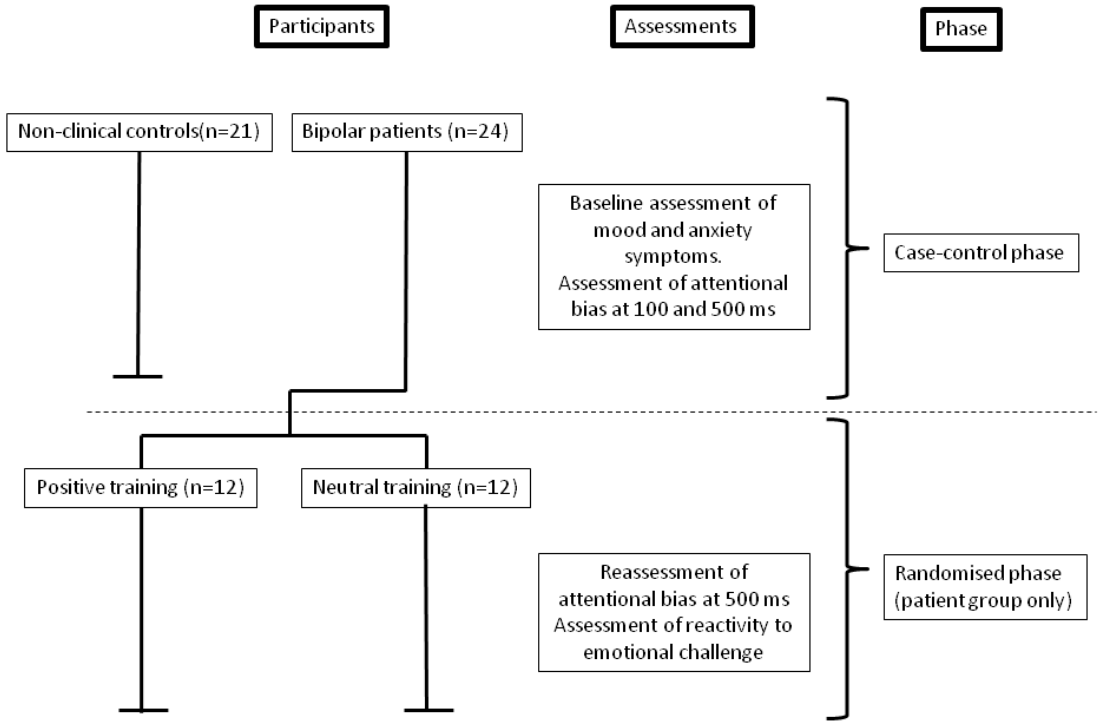



**Figure 6.1: Structure of the study. An initial case-control phase assesses the baseline symptomatology and attentional bias in a group of patients with bipolar disorder and a matched, non-clinical control group. In a second, randomised phase, the patient group are assigned to either positive or neutral attentional training. Following this attentional function is reassessed and reactivity to an emotional challenge is assessed.**

## 6.2.2. Participants

### *6.2.2.1. Patient group*

A total of 25 patients with a diagnosis of either bipolar I or II were recruited via liaison with local psychiatric teams. Patients were provided information on the study by their psychiatric teams if they were older than 18, were not taking regular hypnotic medication [atypical anti-psychotic medications, which are commonly used as a treatment in bipolar disorder (NICE 2006), were permitted], were not sufficiently unwell to warrant in-patient or day patient care and were judged by their clinical team to be able to complete a single, two hour, assessment. Patients who met these criteria were invited to attend the Department of Psychiatry for a

single session which included a clinical assessment and completion of the testing and training tasks. During the clinical assessment the Structured Clinical Interview for the DSM-IV (SCID; Spitzer, Williams et al. 2002) was completed to confirm the diagnosis of bipolar disorder (the diagnosis was confirmed in all patients) and to screen for comorbid axis I diagnoses. One patient who was recruited and randomised was unable to complete the study due to eyesight problems resulting in data being available for 24 patients.

#### *6.2.2.2. Control group*

Participants in the control group (n=21) were recruited via advertisement in the local community. Participants were screened to exclude any current or previous axis I or alcohol misuse disorder using the SCID. Participants were also excluded if they were taking any psychoactive medication (other than hormonal contraceptives) or if they had any familiarity with the tasks or stimuli used during the study. Control participants were also selected on the basis of their age, gender and years of education to ensure that these variables were matched to the bipolar patients at a group level.

### 6.2.3. Questionnaire assessment

#### *6.2.3.1. Self assessment measures*

All participants completed initial self report measures of trait mood (Beck Depression Inventory [BDI]; Beck, Ward et al. 1961), anxiety symptoms (trait subscale of the State-Trait anxiety Inventory [trait-STAI]; Spielberger, Gorsuch et al. 1983), as well as state measures of affect (Positive and Negative Affect Scale [PANAS]; Watson, Clark et al. 1988) and anxiety (state subscale of the STAI [state-STAI]). The state measures were repeated before and after the external stressor task. These were supplemented by a commonly used subjective measure of attentional control (Attentional Control Scale [ACS]; Derryberry and Reed 2002) which

assesses the frequency of 20 different everyday situations which index lapses in attentional control (e.g. “It’s very hard for me to concentrate on a difficult task when there are noises around”). Each situation is rated 0-4 with a lower mark indexing more frequent attentional lapses (min 20, max 80). Participants were also asked to complete a standardised measure of previous suicidality (Beck Scale for Suicide ideation, worst ever [BSS-we]; Beck, Brown et al. 1997) which assess 19 different markers of the severity of suicidal intent (e.g. degree of preparation for suicide attempt) during the period of time identified by the participant as being “the most depressed I have ever been”. Each marker is scored 0-2 with the total score providing an index of suicidal severity (min 0, max 38).

#### *6.2.3.2. Clinician assessment measures*

Standard and well validated observer ratings of depressive (Hamilton Depression Rating Scale [HAM-D]; Hamilton 1960) and manic (Young Mania Rating Scale [YMRS]; Young, Biggs et al. 1978) symptoms were taken for the patient group during the clinical assessment. The HAM-D assesses the severity of 14 different subjective depressive symptoms over the previous 7 days as well as 3 objective measures assessed during the interview itself. Each item is scored and the total score calculated, a lower number indicates lower levels of depressive symptoms (min 0, max 52). The YMRS assesses the subjective and objective extent of 11 symptoms of mania. The total score reflects the severity of manic symptoms with a lower score indicating fewer symptoms (min 0, max 60).

### 6.2.4. Attentional bias training

Attentional bias training was achieved using a word based visual-probe which was largely similar to that employed in Chapter Two. This involved the presentation of word pairs consisting of negative and neutral words for 500 ms. The specific word pairs presented in the

training task included both the negative-neutral pairs used in Chapter Two (taken from; MacLeod, Rutherford et al. 2002) as well as 124 extra negative-neutral word pairs which were drawn from a range of published sources (Hope, Rapee et al. 1990; Mattia, Heimberg et al. 1993; Asmundson and Stein 1994; Boyer, Compas et al. 2006; Munafo, Hayward et al. 2006; Murphy 2007). The rationale for expanding the number of word pairs used was that the original set of words contained predominantly physically rather than socially threatening negative words. Previous work has demonstrated that attentional biases in clinical disorders tend to be more evident when the stimuli used in the assessment tasks are tailored to the specific concerns of the patient group (Mathews and MacLeod 1985; Becker, Rinck et al. 2001). It was expected that socially threatening stimuli would be more closely linked to anxiety than physically threatening stimuli for patients with bipolar disorder as they are more likely to describe the previous negative emotional experiences commonly encountered by patients (e.g. "die" is an example of a physically threatening word whereas "cry" is an example of a socially threatening word). As it was expected that a clinical group would find extended training more difficult to complete, a shorter training paradigm with two booster sessions was administered; the initial training consisted of 220 trials with a single rest session, booster sessions contained 110 trials with no rest session (in Chapter Two the initial session consisted of 576 trials and the single booster session consisted of 144 trials). Lastly, as the current study focussed on the possible therapeutic potential of training, the two training conditions used were; a positive training regime (probes always replace the neutral word) and a neutral regime (probes randomly replace either word; see Chapter Five). Word pairs were divided into two sets with one set being used for all training sessions and the other set being used for the testing sessions.

## 6.2.5. Assessment of attentional bias

### *6.2.5.1. Word visual-probe*

The word visual-probe presented the set of word stimuli not used in training. Baseline assessment included a block of 48 trials in which the stimuli were presented for 100 ms and a block in which they were presented for 500 ms. These relatively brief durations were selected as they have previously been associated with symptoms of anxiety (see Chapter One). After training attention was assessed at 500 ms only (NB the same presentation time as training) in order to reduce the duration of the testing session.

### *6.2.5.2. Face visual-probe*

Face stimuli from a range of sources (Ekman and Friesen 1976; Lundqvist, Flykt et al. 1998; Tottenham, Tanaka et al. 2009) displaying neutral faces were paired with faces displaying either anger or fearful faces. Images subtended a vertical visual angle of ≈11° and were separated by ≈3.5°. The timings and number of trials used in the face visual-probe task were identical to those used in the word task.

## 6.2.6. Emotional challenge task

The emotional challenge task used was the social failure task described in Chapter Five. The selection of this task was again motivated by the predication that anxiety in bipolar disorder would be closely linked to the socially threatening dimension of situations (see Section 6.2.4). During the task participants are required to identify a fourth word which is semantically linked to three other words which are presented on the screen. They are informed that the task measures "creative intelligence" and, following the task, their responses are marked and their overall score is fed back to them by the experimenter. The task is extremely difficult resulting in participants performing more poorly than expected and

causing a reliable increase in negative emotions (McFarlin and Blascovich 1984; Moberly and Watkins 2006). Importantly, the increase in negative symptoms experienced by non-clinical participants in response to this task was found to be influenced, at a trend level, by attentional bias training in Chapter Five. State measures of mood (negative-PANAS) and anxiety (state-STAI) were collected before and after completion of the task.

### 6.2.7. Statistical analysis

Chi-squared or t-tests were used to assess baseline demographic differences between groups. Consistent with previous chapters, median reaction times of accurate trials were used in the analyses. These were then entered into mixed model ANOVAs to test both baseline differences and the effects of training within the patient group. For the baseline assessment a between subject factor of clinical status (non-clinical control, patient) and a within subject factors of stimuli duration (100 ms, 500 ms) and probe location (probe replaces negative stimuli, probe replaces neutral stimuli) was used. The effect of anxiety on baseline biases was assessed both by restricting the analysis to clinical participants with a comorbid anxiety disorder (Brotman, Rich et al. 2007) and by correlating vigilance scores with subjective anxiety symptoms (measured using the trait-STAI) in the clinical group. In order to assess the effects of training, a mixed model ANOVA with a between subject factor of attentional training group (positive training, neutral training) and within subject factors of time (before training, after training) and probe location (probe replaces negative stimuli, probe replaces neutral stimuli) was used. Changes in self report mood and anxiety across the challenge task were assessed in the same fashion as Chapter Five; the difference scores for the relevant measure were entered into a univariate ANCOVA which included training group (positive,

neutral) as a between subject factor and the baseline mood or anxiety score entered as a covariate.

A supplementary analysis of the data was performed to assess the influence of medication status in the patient group. A "medication load" score was calculated for each participant following the procedure outlined by Almeida and colleagues (Almeida, Versace et al. 2010). The score is derived by assigning a value (0=absent, 1=low dose, 2=high dose) to each psychotropic medication a patient is taking, with the total score for the patient reflecting their overall "medication load". The medication load score was correlated against baseline questionnaire and vigilance scores for the patient group and added as a covariate when analysing the effects of training. As in Chapter Two, the vigilance score was calculated for the visual probe tasks such that a more positive value indicated greater attentional bias towards the threatening stimuli.

## 6.3. Results

### 6.3.1. Demographic data and baseline anxiety

Table 6.1 displays the baseline demographic and questionnaire measurements from all groups. As predicted, the patient group reported significantly more anxious and depressive symptoms than the control group. Further, mean scores of the self report scales of trait anxiety (trait-STAI), depression (BDI) and previous suicidality (BSS-we) as well as the clinical rating of depressive (HAM-D) and manic (YMRS) symptoms were comparable with previous studies of euthymic bipolar patients (Clark, Iversen et al. 2002; Jones, Scott et al. 2005; Simon, Zalta et al. 2007; Chandler, Wakeley et al. 2009). Interestingly, the patient group reported significantly lower levels of attentional control (ACS), a previously

unreported subjective effect. A significant negative correlation was found between trait anxiety (trait-STAI) and attentional control (ACS) in the patient group [$r(24) = -0.53$, $p = 0.008$] suggesting that the increased levels of anxiety found in the patient group may be partially attributable to difficulties in attentional control (Eysenck, Derakshan et al. 2007). None of the baseline variables were significantly correlated with medication load within the patient group [all $p > 0.28$].

The randomisation procedure within the patient group was largely successful in that the positive and neutral training groups were well matched for most baseline characteristics. The single exception to this was for years of education which was greater for the positive than neutral training group.

### 6.3.2. Patient characteristics

The symptomatology and medication status of the patient group are summarised in Table 6.2. As can be seen, the majority of patients had a diagnosis of Bipolar I disorder, there was a high rate of comorbidity and the majority of patients were medicated. 15 of the 24 patients (62%) had a comorbid anxiety disorder.

**Table 6.1: Demographic details and baseline questionnaire measures from patient and control groups.**

| Measure | Patient groups | | *p* value for difference between training groups | Control group (n=21) | *p* value for difference between control and patient groups |
|---|---|---|---|---|---|
| | Positive Training (n=12) | Neutral Training (n=12) | | | |
| **Age** | 40 (13.4) | 37.3 (14.8) | 0.65 | 33.8 (11.3) | 0.21 |
| **Gender** | 6:6 | 6:6 | 1.0 | 11:10 | 0.87 |

| | | | | | |
|---|---|---|---|---|---|
| **(F:M)** | | | | | |
| **Years of education** | 16.1 (2.3) | 13.5 (3.2) | 0.04 | 15.6 (3) | 0.42 |
| **HAM-D** | 5.8 (4) | 4.4 (4.8) | 0.44 | - | - |
| **YMRS** | 2.2 (2.2) | 2.5 (3.1) | 0.76 | - | - |
| **BDI** | 13.3 (8.6) | 11.8 (9.8) | 0.68 | 4.6 (4.5) | 0.001 |
| **trait-STAI** | 47 (18.3) | 47.5 (10.9) | 0.94 | 32.9 (47.3) | <0.001 |
| **ACS** | 46.9 (11.1) | 49.9 (7) | 0.44 | 59.5 (8.4) | <0.001 |
| **BSS-we** | 14.1 (12) | 14.7 (12) | 0.91 | 2 (3.7) | <0.001 |
| **pos-PANAS** | 29.3 (9.3) | 29.9 (7.1) | 0.86 | 34.6 (7.4) | 0.03 |
| **neg-PANAS** | 14.2 (4.7) | 12.2 (3.4) | 0.25 | 11.6 (3.8) | 0.2 |
| **state-STAI** | 35.5 (9.3) | 33.6 (7.8) | 0.59 | 28.6 (9) | 0.03 |

HAM-D= Hamilton Depression Rating Scale; YMRS=Young Mania Rating Scale; BDI=Beck Depression Inventory; STAI=State-Trait Anxiety Inventory; ACS=Attentional Control Scale; BSS-we=Beck Scale for Suicide Ideation- worst ever; PANAS=Positive and Negative Affect Scale.

**Table 6.2: Clinical summary of patients**

| Number | Gender | Age | Bipolar Diagnosis (Bipolar I or II)[a] | HAM-D | YMRS | Time (years) and type of last mood episode | Comorbid Diagnosis | Psychotropic Medication (total daily dose) | Medication Load[b] |
|---|---|---|---|---|---|---|---|---|---|
| 1 | M | 63 | I | 0 | 1 | 2.5 Mania | Nil | Carbamazepine 800mg | 1 |
| 2 | F | 44 | I | 8 | 2 | 9 Depression | Panic disorder and agoraphobia | Lithium 200mg. Valproate 125mg | 2 |
| 3 | F | 53 | II | 2 | 0 | 3 Depression | Social phobia | Sertraline 200mg. Amisulpiride 50mg | 3 |
| 4 | M | 46 | I | 1 | 1 | 0.17 Mania | Nil | Carbamazepine 1000mg. Risperidone 1mg. Lamotrigine 1000mg | 5 |
| 5 | F | 60 | II | 9 | 0 | 1 Depression | Nil | Nil | 0 |
| 6 | M | 40 | II | 12 | 1 | 8 Depression | Hypochondrias | Valproate 1.5g. Lamotrigine 300mg. Citalopram 30mg. | 6 |
| 7 | M | 27 | I | 0 | 0 | 3 Depression | Harmful drug use (previous) | Lithium 400mg. Melatonin 6mg. Risperidone 4mg. Valproate 750mg | 4 |
| 8 | M | 37 | I | 5 | 1 | 2 Depression | Generalised anxiety disorder. Drug dependency (currently in remission) | Valproate 1000mg | 2 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 9 | F | 21 | II | 4 | 5 | 0.17 Hypomania | Bulimia | Nil | 0 |
| 10 | F | 21 | I | 2 | 1 | 0.25 Depression | Panic disorder without agoraphobia. Harmful alcohol use (current) | Valproate 300mg | 1 |
| 11 | M | 44 | I | 15 | 2 | 0.75 Depression | Generalised anxiety disorder | Citalopram 20mg, Valproate 1.2g | 2 |
| 12 | F | 22 | I | 3 | 3 | 1 Mania | Panic disorder with agoraphobia | Nil | 0 |
| 13 | M | 32 | I | 7 | 0 | 0.17 Depression | Generalised anxiety disorder | Lithium 1g, Amisulpiride 800mg. Olanzapine 15mg. | 6 |
| 14 | M | 51 | I | 8 | 0 | 0.04 Mania | Nil | Quetiapine 50mg | 1 |
| 15 | M | 39 | II | 3 | 11 | 1 Depression | Harmful alcohol use (previous) | Quetiapine 300mg, Valproate 500mg | 3 |
| 16 | F | 27 | I | 2 | 3 | 0.67 Depression | Harmful alcohol use (previous) | Quetiapine 200mg, Lamotrigine (on trial so may not be taking) | 1 |
| 17 | F | 36 | I | 3 | 7 | 6 Depression | Harmful drug use (previous). Panic disorder with agoraphobia. OCD. | Fluoxetine 20mg, Quetiapine 300mg, Valproate 1.5g | 5 |
| 18 | M | 25 | II | 12 | 3 | 0.17 Depression | Panic disorder with agoraphobia, OCD, generalised anxiety disorder | Citalopram 10mg | 1 |
| 19 | M | 32 | I | 1 | 3 | 0.33 Depression | Panic disorder with agoraphobia, OCD, alcohol dependence (currently in remission). Harmful drug use (previous) | Lithium 1.2g, Quetiapine 50mg | 3 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 20 | F | 25 | I | 11 | 5 | 0.67 Mania | Alcohol and drug dependence (currently in remission). Generalised anxiety disorder. Panic disorder with agoraphobia | Nil | 0 |
| 21 | F | 19 | I | 2 | 1 | 0.67 Depression | Panic disorder with agoraphobia. Bulimia. | Aripiprazole 20mg, Fluoxetine 10mg, Lamotrigine 200mg | 5 |
| 22 | F | 53 | I | 0 | 0 | 8 Depression | Anorexia | Olanzapine 1.25mg | 1 |
| 23 | M | 48 | I | 7 | 3 | 1.5 Mania | Panic disorder and agoraphobia | Valproate 1.25g | 2 |
| 24 | F | 63 | II | 6 | 3 | 0.17 Depression | Panic disorder and agoraphobia | Lithium 600mg, Fluoxetine 60mg | 3 |

[a] Using DSM-IV criteria, a diagnosis of bipolar I disorder requires episodes of mania, which significantly impact on an individual's functioning whereas a diagnosis of bipolar II disorder is made if episodes of depression occur interspersed with more mild episodes of hypomania. [b] Medication Load provides a summary statistic incorporating the number and relative dosages of psychotropic medication (Almeida, Versace et al. 2010). OCD=Obsessive Compulsive Disorder

### 6.3.3. Baseline attentional biases

Analysis of the baseline data from the faces visual-probe task (Figure 6.2a and Table 6.3) revealed a significant group x stimulus duration x probe location interaction [$F(1,43) = 4.6$, $p = 0.04$], indicating that the groups displayed different baseline attentional biases to the threatening faces, but that this depended on the duration of stimuli presentation. Analysing the 100 ms and 500 ms trials separately revealed that, while numerically the patient group showed a greater attentional bias towards threat at 500 ms and a smaller bias at 100 ms, neither effect was significant [100 ms; $F(1,43) = 1.6$, $p = 0.21$. 500 ms; $F(1,43) = 2.4$, $p = 0.13$]. There was no evidence for baseline differences in attentional bias during the words task [$F(1,43) = 0.48$, $p = 0.49$]. These results did not change when analysis was restricted to the subgroup of patients with a comorbid anxiety disorder (c.f.; Brotman, Rich et al. 2007) and, within the patient group, levels of anxiety as measured by the trait subscale of the STAI did not covary with mean attentional bias [all $p$'s $> 0.26$]. However, analysis of both tasks also revealed that the patient group was significantly slower on average when responding than the control group, regardless of the trial type [words; $F(1,43) = 4.3$, $p = 0.04$. faces; $F(1,43) = 5.5$, $p = 0.02$].

In order to further investigate the influence of the reaction time differences between groups, an index of average reaction time was calculated by taking the average reaction time from all four trial types (see Table 6.3) for each participant, separately for the faces and words tasks. This variable was then used in subsequent analyses of the vigilance scores from the visual probe tasks. Entering this variable as a covariate into the analysis of the faces task did not change the pattern of results reported above, with the exception that the difference between the groups at 100 ms was rendered significant [greater attentional bias to threatening faces in

the control group at 100 ms; $F(1,42) = 7.1$, $p = 0.01$]. Additionally however, across both groups and both stimulus durations, participants' average reaction time was found to significantly covary with their vigilance towards the threatening faces [$F(1,42) = 8.8$, $p = 0.005$] such that the longer the average reaction time, the greater the attentional bias to threat at both durations (Figure 6.2b). The relationship between vigilance and average reaction time was independently apparent in the patient group [$r(24) = 0.43$, $p = 0.04$] and present at a trend level in the control group [$r(21) = 0.40$, $p = 0.07$][4]. Adding average reaction time as a covariate for the words task produced no significant changes to the analyses.

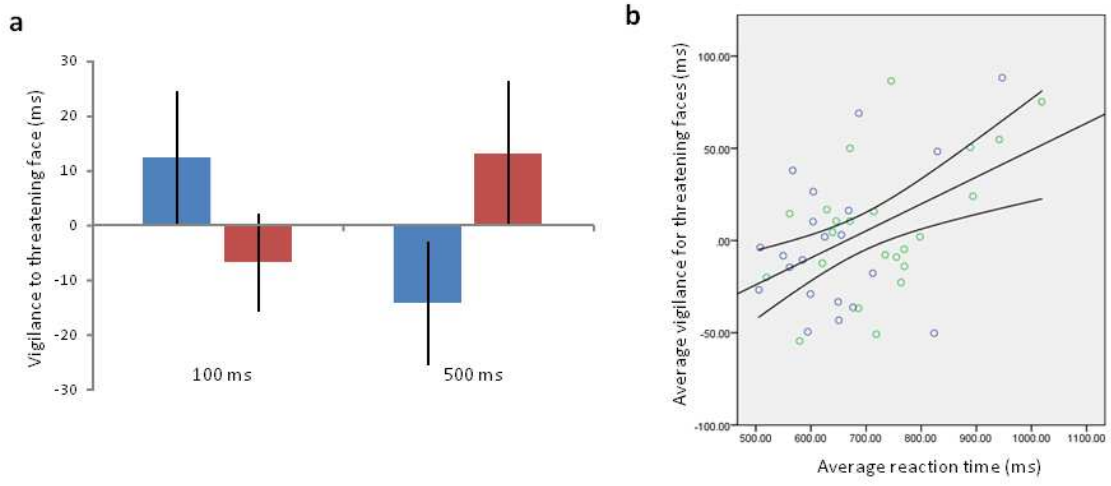



**Figure 6.2: Data from the faces visual-probe task. a) baseline differences in attentional bias in the patient group (red bars) and control group (blue bars). Although the overall group x stimulus duration interaction was significant, there was no significant effect of group at either 100 ms or 500 ms. Bars represent mean (SEM) of vigilance scores. b) Positive correlation between average reaction time on all trials and the estimated vigilance for threatening faces. Patient group = green circles, control group = blue circles. A regression line (incorporating all participants) has been superimposed.**

## 6.3.4. Effect of training on attentional bias within the patient group

[4] This unexpected finding, of a trend positive correlation between average reaction time and vigilance score in the control group, prompted a reanalysis of the behavioural data from earlier chapters (Two and Five) to assess whether a similar relationship between these variables was apparent in these studies. However, no significant correlation was found between average reaction time and attentional bias in any of the tasks reported in these studies.

There was no evidence that training altered attentional bias (Table 6.3). This was true when change in bias across training was considered [words; $F(1,22) = 0.96$, $p = 0.32$. faces; $F(1,22) = 0.87$, $p = 0.36$] and when restricting analysis to the post training assessment, as was done in Chapter Two [words; $t(22) = 0.65$, $p = 0.52$. faces; $t(22) = 1.15$, $p = 0.26$]. Adding the number of years of education completed by each participant (NB as described in section 6.3.1 the training groups differed on this measure) or the average reaction time (which did not differ between the groups; $t(22) = 0.15$, $p = 0.88$) as covariates did not alter these results. The results were also not changed by restricting analyses to those patients with a comorbid anxiety diagnosis.

### 6.3.5. Effect of training on emotional reactivity within the patient group

The emotional challenge task (Table 6.4) caused an increase in the state-STAI scores [$F(1,21) = 9.7$, $p = 0.005$] but not the neg-PANAS [$F(1,21) = 0.1$, $p = 0.75$]. The training completed did not modify either of these effects [state-STAI; $F(1,21) = 0.001$, $p = 0.98$. neg-PANAS; $F(1,21) = 0.5$, $p = 0.48$]. Adding years of education or average reaction time as covariates or restricting analyses to patient with comorbid anxiety did not alter these findings.

### 6.3.6. Effects of medication

The psychotropic medication taken by patients, as measured by the "medication load" (see section 6.2.7), had little effect on any of the outcome measures. Specifically, there was no correlation between medication load and baseline vigilance scores within the patient group for either the words or faces task [all *p's* > 0.11]. Medication load also did not significantly impact on the effects of training on either the attention tasks [all *p's* > 0.4] or emotional

reactivity [all *p's* > 0.15]. There was a weak trend for medication load to positively correlate with average reaction time [$r(24) = 0.34$, $p = 0.1$].

**Table 6.3: Raw group reaction time data from visual-probe tasks. All reaction times are displayed as mean (standard deviation) to the nearest ms.**

| Time | Before Training | | | | | | After Training | |
|---|---|---|---|---|---|---|---|---|
| **Group** | **Control** | | **Positive Training** | | **Neutral Training** | | **Positive Training** | **Neutral Training** |
| **Stimuli Duration (ms)** | **100** | **500** | **100** | **500** | **100** | **500** | **500** | **500** |
| **Neutral Words** | 682 (106) | 692 (154) | 771 (167) | 767 (124) | 767 (164) | 772 (135) | 671 (107) | 680 (119) |
| **Negative Words** | 699 (100) | 699 (154) | 769 (133) | 780 (132) | 781 (166) | 761 (113) | 654 (98) | 681 (114) |
| **Neutral Faces** | 656 (132) | 638 (115) | 718 (137) | 748 (158) | 716 (125) | 737 (134) | 667 (119) | 654 (127) |
| **Negative Faces** | 644 (105) | 653 (114) | 722 (122) | 740 (138) | 726 (108) | 719 (124) | 668 (109) | 676 (118) |

**Table 6.4: Raw group data from the social failure task. Numbers indicate the mean (standard deviation) from the relevant scales.**

| Group | Positive Training | | Neutral Training | |
|---|---|---|---|---|
| **Time** | **Before Stressor** | **After Stressor** | **Before Stressor** | **After Stressor** |
| **Negative-PANAS** | 13.7 (4.0) | 19.1 (9.9) | 14.5 (6.0) | 18.2 (8.0) |
| **State-STAI** | 37.8 (12.1) | 47.4 (12.9) | 39.1 (11.9) | 48.2 (11.9) |

Negative-PANAS= negative subscale of the Positive and Negative Affect Scale. State-STAI = state subscale of the Spielberger State-Trait Anxiety Inventory

## 6.4. Discussion

In the current chapter attentional function and its malleability were assessed in a group of patients with bipolar disorder. While the predicted baseline differences in anxious symptoms between the patient and control groups were demonstrated, the predicted differences in attentional bias were not. Further, the single session of attention training, completed by the patient group, did not influence either the measures of attentional function or emotional reactivity. Most importantly, however, the current analysis suggested that the estimate of attentional bias derived from the visual-probe task was contaminated by significant average reaction time differences between the clinical and control groups. These results suggest that this commonly used behavioural measure of attentional bias may be unreliable when comparing clinical and control groups which differ on average reaction time.

### 6.4.1. Baseline attentional bias

As reviewed in the introduction, previous attempts to characterise attentional bias in bipolar disorder have produced mixed results (Bentall and Thompson 1990; Lyon, Startup et al. 1999; Kerr, Scott et al. 2005; Brotman, Rich et al. 2007; Jongen, Smulders et al. 2007). However, while there has been little consistency in the attentional biases reported in bipolar disorder, one finding has been replicated; in all patient studies in which the relevant analysis was reported (Lyon, Startup et al. 1999; Kerr, Scott et al. 2005; Jongen, Smulders et al. 2007) the patient group was found to have a significantly slower average reaction time than the control group. The current study replicates this finding and, additionally, demonstrates that the *average* time participants took to respond significantly predicted their estimated attentional vigilance to the threatening faces. The critical importance of this finding is that it violates the assumptions underlying the interpretation of the visual probe task. Specifically,

the task estimates the spatial deployment of attention as the relative reaction times of trials in which the probe replaces the emotional stimuli compared to those when it replace the neutral stimuli (i.e. the vigilance score, see Chapter One; MacLeod, Mathews et al. 1986). Relative reaction time scores are used to account for inter-individual differences in gross reaction time which are considered to be unrelated to the spatial deployment of participants' attention. This assumption predicts that vigilance scores and mean reaction times should be unrelated and is plainly violated by the current results where the two measures were found to strongly covary. While it is possible to statistically control for differences in average reaction time, as done in the current study, the very finding that increased average reaction time is associated with increased vigilance towards threatening information implies the existence of some extra process, which influences the vigilance score and is not modelled by the task design, and thus raises doubts about how best to interpret the data. In summary, the demonstrated relationship between average reaction time and vigilance score in the faces visual-probe task suggests that the estimates of attentional bias derived from the task may be unreliable.

The significant correlation between average reaction time and attentional bias must therefore cast doubt on the interpretation of the difference in attentional deployment identified between the patient and control groups in the current study; an interaction between group and stimulus duration was found such that the control group demonstrated greater attentional bias to threatening faces at 100 ms and less at 500 ms. Given the concerns with interpretation of these results described above, it is unclear whether this unpredicted finding accurately represents the difference in attentional function between the two groups. Alternative measures of attentional deployment, which do not depend on reaction time outcomes, such as

that obtained from eye-tracking technologies, may provide a more reliable assay of attentional function in this group of patients in future studies.

Given the frequency with which increased reaction times have been reported in previous studies of bipolar patients, the current results raise doubts about the interpretation of the attentional bias data from those studies which have utilised the visual probe task without accounting for gross reaction time differences (Brotman, Rich et al. 2007; Jongen, Smulders et al. 2007). Indeed, the results may also be relevant to interpretation of emotional Stroop data (Lyon, Startup et al. 1999; Kerr, Scott et al. 2005), which also relies on reaction time differences to infer attentional bias although, conceptually at least, this task measures a distinct attentional effect (response interference as opposed to selective spatial attention; see Chapter One) and therefore may be less affected.

The current study is unable to specify the mechanism by which average reaction time is associated with the estimates of threat vigilance generated by the visual probe task. The trend level positive correlation between medication load and average reaction time in the patient group raises the possibility that some of the variance in reaction time may be attributed to the effects of medication. However, even if one assumes the significance of this finding, it does not account for the observed relationship between reaction time and vigilance. One interesting, if speculative, possibility is that participants' average reaction time indexes an interference effect generated by the context of the experiment, such that the more participants' attention is captured by the experimental context, the slower they are in performing the task itself. This would suggest that the correlation between the visual probe derived vigilance score and average reaction time demonstrated in the current study arises as a consequence of these measures assessing similar attentional effects; the spatial deployment

of attention towards threatening stimuli and general task interference induced by the (presumably threatening) task context. This possibility may be explored in future studies by employing additional measures which are thought to be sensitive to the threatening effects of task context, such as the emotional startle (Grillon and Baas 2003). However, a number of other explanations may account for this finding and, as indicated above, the current study is unable to adjudicate between these.

### 6.4.2. Effects of attentional training

As the effect of attentional bias training was assessed using within subject changes in vigilance score, this analysis was less susceptible to the confounding effects of average reaction time (which did not differ between the training groups). Despite this no effect of training was found on either the measures of attentional deployment or emotional reactivity. A number of differences between the methodologies used in the current study and those employed in previous chapters may account for the failure to replicate the basic training effect. Most obviously, recruitment of a sample of patients with significant mental ill health and well documented difficulties with sustained attention (Torres, Boudreau et al. 2007) may have decreased the impact of a single session of training. Alternatively, alterations to the training regime used in the current study may be responsible. The training regime employed was significantly shorter (440 vs. 720 trials) than that used in Chapters Two and Three and used neutral (rather than negative) training as the control condition. Both of these factors may have reduced the size of any training effect and have lead to the observed negative results. The decision to use a truncated training regime was prompted by concerns that the patient group would find a longer session difficult, and thus future studies assessing the impact of attentional training in bipolar disorder may benefit from using multiple, shorter training

episodes as was done in Chapter Five. An associated methodological limitation to the current study was that the control group did not complete the training task. Had they done so, discrimination between the various patient-based and task-based explanations for the negative results described above may have been facilitated as demonstration that training was effective in the control group would give confidence that the training task itself worked as expected.

### 6.4.3. Attentional control in bipolar disorder

Analysis of the baseline self report measures in the patient and control groups revealed increased difficulties with attentional control, as measured by the ACS, in the patient group. Although analysis of these data was not a primary goal of the current study, difficulties with attentional control have previously been proposed as a causal mechanism underlying anxiety (Eysenck, Derakshan et al. 2007) and therefore the low levels of this measure found in the patient group may account for some of the anxiety symptoms reported by this group. Consistent with this proposal a significant negative correlation was found between trait anxiety (trait-STAI) and attentional control (ACS) in the patient group. Characterisation of the relationship between attentional control and anxiety symptoms in bipolar patients may be enhanced by the use of non-self report measures of attentional control, such as the antisaccade task (Derakshan, Salt et al. 2009), in future studies.

### 6.4.4. Limitations

The current study has a number of further limitations which should be acknowledged. Firstly, the sample sizes used were relatively small. While a patient group of 24 compares favourably with other published studies (Bentall and Thompson 1990; Kerr, Scott et al. 2005; Brotman, Rich et al. 2007) it may be that a larger sample size would have led to the detection of significant differences between the patient and control groups. Secondly, patients with

comorbid diagnoses and a range of different medications were recruited. It is conceivable that stricter recruitment criteria would have resulted in a more homogenous patient sample and a more accurate estimate of attentional function. However, the use of strict inclusion criteria tends to also reduce the clinical validity of studies by excluding many of the patients who are routinely encountered in clinical practice; for example, the majority of patients with bipolar disorder would be excluded from most published treatment trials of the disorder (Storosum, Fouwels et al. 2004) raising questions about whether the results of the studies can be generalised to the patients seen in practice. As the current study sought to assess the influence of anxiety on bipolar disorder, excluding comorbid diagnoses was inappropriate although excluding patients who were taking the more sedative antipsychotic medications (e.g. Quetiapine, Olanzapine) or those more likely to cause Parkinsonism (e.g. Risperidone) may have improved the sensitivity of the cognitive measures by reducing patient sedation and improving motor function respectively.

### 6.4.5. Conclusion

In conclusion, the current study was unable to find convincing evidence for biased attentional function using a visual-probe task in a clinically representative group of patients with bipolar disorder and was unable to demonstrate any effect of a single episode of attentional training. Importantly, analysis indicated that differences in average reaction time between the clinical and control groups significantly impacted the estimates of attentional function derived from a visual-probe task. These findings cast doubt on the interpretation of previous studies which have employed similar methodologies in this clinical group.

# Chapter 7

# General Discussion

## 7.1. Summary of findings

This thesis aimed to investigate the mechanisms by which attentional bias training alters the habitual deployment of attention. The thesis then sought to use these initial results to improve the understanding of attentional control processes and, ultimately, to guide the development of novel treatment strategies. In order to do this a series of experimental studies were completed which characterised the behavioural and neural effects of training, explored the nature of the attentional control processes elicited by training, assessed the impact of training when combined with antidepressant medication and finally, assessed the possible utility of training in patients with bipolar disorder.

In the first experiments (Chapter Two), the behavioural effects of word and faced based training were assessed in a non-clinical population. These studies confirmed that the training procedures employed in the thesis produced the expected basic effect; alteration of attentional deployment to subsequently presented affective stimuli. Crucially, however, it was also found that these effects generalised to a novel type of stimuli not used in the training task, with word based training leading to an alteration of attention to emotional faces (Figure 7.1). These results suggest that training regimes can alter attentional function to emotional stimuli generally (see also, Chapter Five and section 7.2.1), and that they therefore may be able to influence the general attentional biases found in the emotional disorders (MacLeod and

Mathews 1988; Yiend and Mathews 2001; Gotlib, Krasnoperova et al. 2004; Fox, Russo et al. 2005; Mathews and MacLeod 2005; Joormann, Talbot et al. 2007).

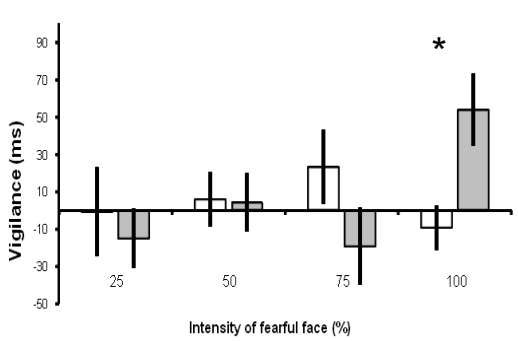



**Figure 7.1: Word based training influences attentional deployment to novel faces, taken from Chapter Two. Vigilance towards emotional faces presented for 500 ms. Vigilance is calculated such that a positive number indicates an attentional bias towards the negative stimulus, and a negative number indicating an attentional bias away from the negative stimulus. The intensity of the morphed negative face is arranged along the x axis. White bars = avoid-negative training, which induces a benign attentional bias. grey bars = attend-negative training, which induces a negative bias. * = p≤0.05.**

The results from Chapter Two were used as the basis for a subsequent investigation of the neural effects of training (Chapter Three). In this study, attentional bias training was found, as predicted (DeRubeis, Siegle et al. 2008; Harmer 2008), to tune the activity of prefrontal control systems in response to emotionally valenced facial stimuli. Specifically, following a single episode of either avoid-negative or attend-negative attentional training, activity in the lateral prefrontal (lPFC) and rostral anterior cingulate (rACC) cortices increased when participants were required to violate the rules which had governed their training (Figure 7.2). Further, the observed connectivity between the lPFC control areas and face specific visual association cortex was consistent with the frontal areas acting to control attention to the stimuli. Comparison of these results with previous studies which demonstrated an effect of antidepressant medication on amygdala function (Sheline, Barch et al. 2001; Fu, Williams et al. 2004; Harmer, Mackay et al. 2006; Murphy, Norbury et al. 2009) supports the proposal made in Chapter One that pharmacological and cognitive manipulations target different attentional control systems.

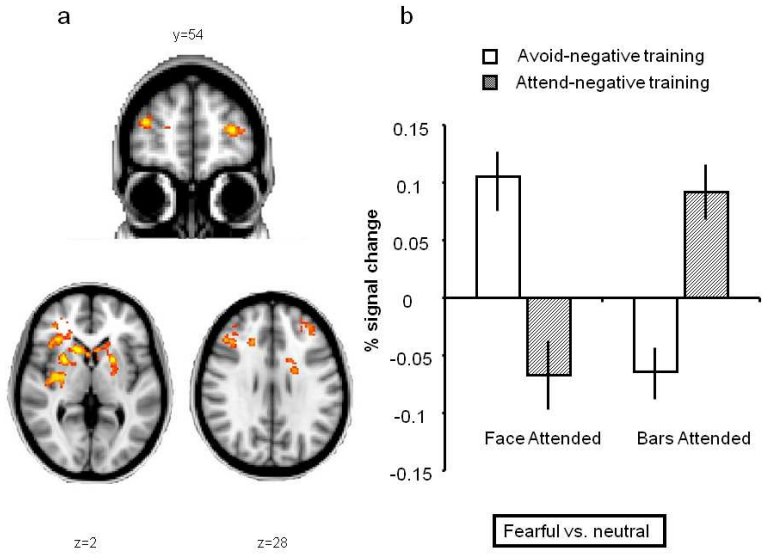



**Figure 7.2: Effect of attentional bias training on BOLD signal, taken from Chapter Three. (a) Whole brain, cluster corrected (*Z*-threshold = 2.3, $p < 0.05$) analysis demonstrating bilateral frontal and striatal regions in which activity corresponded to the effects of attentional bias training on the emotion x attention interaction. (b) The mean (SEM) percent signal change associated with the fear vs. neutral face contrast extracted from the right lateral PFC cluster (other clusters show an identical pattern). Estimates for the fearful face-neutral face contrast are displayed separately for trials in which participants had to attend to the location of the face (face attended) or to the location of the bars (bars attended). In all clusters activation is greatest when participants direct their attention contrary to their training; thus the avoid-negative training group (white bars), who have been trained to look away from negative and towards neutral stimuli, show increased activation when looking towards negative and away from neutral stimuli. The attend-negative training group (gray bars) show the opposite pattern of activation.**

It was suggested that the activity in the frontal control areas elicited by attentional bias training was similar to the prediction error signals described in reinforcement learning paradigms (Sutton and Barto 1998). This observation prompted a novel analysis of the imaging data (Chapter Four) with the aim of assessing whether an expectancy based computational model could predict the trial-by-trial alterations of attention to emotional stimuli. Two key model derived parameters, threat expectancy and surprise (prediction error magnitude), significantly predicted both behavioural and neural indices of attention to the emotional stimuli (Figure 7.3). By demonstrating that participants continuously track threat-expectancy and use it to influence the deployment of attention these results confirmed the novel proposal that expectancy based processes are involved in the control of attention to emotional stimuli.

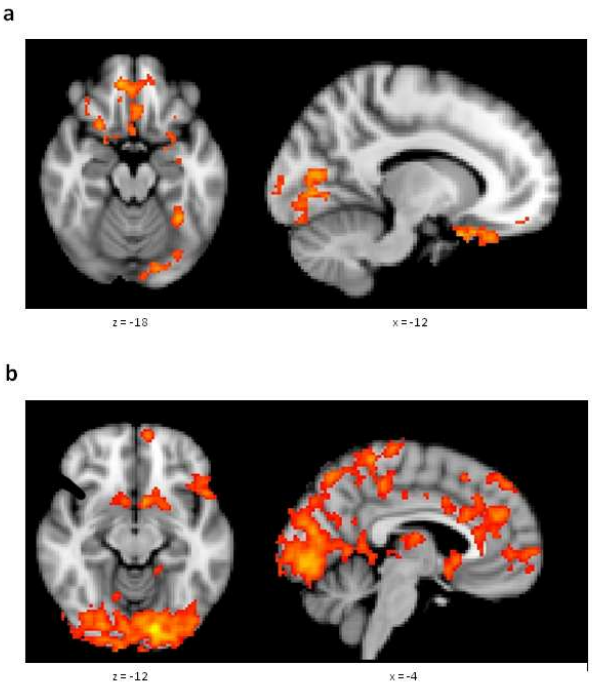



**Figure 7.3: Results of whole brain, cluster corrected (*Z*-threshold = 2.3, $p < 0.05$) analyses, taken from Chapter Four, showing activity associated with a) the threat-expectation and b) surprise regressors. a) threat-expectation was found to covary with activity in the left ventral visual stream and with attentional control regions including the OFC and left amygdala. b) surprise was associated with activation of the visual cortex and a network of frontal regions including the ACC, dlPFC and ventral striatum (not shown). Average activations maps across all participants have been rendered over the standard MNI brain.**

The final two data chapters of the thesis were designed to address clinically relevant questions relating to the use of attentional bias training. The first study (Chapter Five) investigated the effect of combining attentional training with antidepressant medication on a range of behavioural measures of emotional processing and on emotional reactivity to two forms of external challenge. The results of the study suggest that the interventions interfered with each other at a cognitive level, with combined administration producing a smaller effect on emotional memory than either intervention on its own (Figure 7.4).

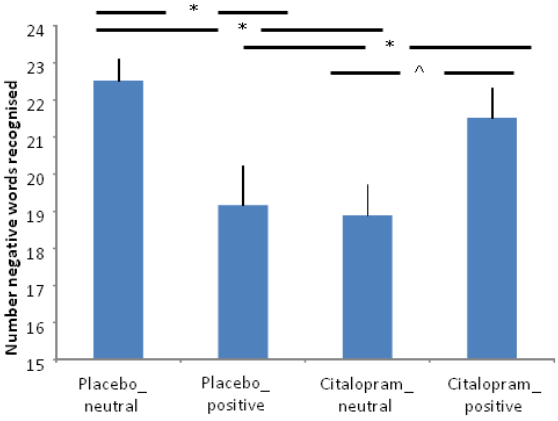



**Figure 7.4: Effects of antidepressant medication, attentional training and their interaction on recognition of negative words, taken from Chapter Five. Both citalopram and training significantly decrease recognition of negative words whereas the combination treatment has no effect. All graphs display mean (SEM). * = p < 0.05, ^ = p < 0.1.**

However, this interference effect was not evident in the emotional reactivity of participants; in this case the two treatments were found to have distinct and specific effects, with each one protecting against the negative mood induced by a particular emotional challenge. These results raise the interesting possibility that the emotional impact of interventions which modify emotional biases (such as antidepressant medication or attentional bias training) may be contingent on the particular form of external stress subsequently experienced. More generally however, these results demonstrate that healthy volunteer models (Harmer, Hill et al. 2003; Harmer, Shelley et al. 2004; Browning, Reid et al. 2007; Murphy, Yiend et al. 2009) may be used to examine complex, but clinically relevant treatment interactions.

The final study of the thesis assessed attentional function and it's malleability in a clinically representative sample of patients with bipolar disorder (Chapter Six). The study was unable to identify a convincing baseline attentional bias in the clinical group, despite that group showing significantly raised levels of anxiety, and was unable to detect any effect of the attentional training regime. Further analysis (Figure 7.5) of the results suggested that the estimates of attentional function derived from visual-probe tasks may be unreliable when comparing groups, such as the patient and control groups in the current study, with grossly different reaction times. This suggests that tasks which do not rely on speeded reaction times may provide a more reliable measure of attentional function in certain clinical groups.

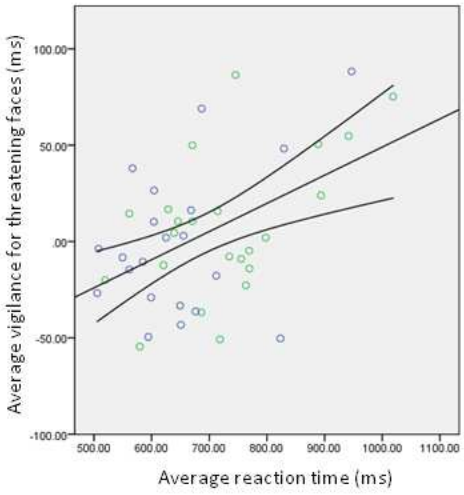



**Figure 7.5: Significant positive correlation between average reaction time and attentional vigilance taken from the faces visual-probe task in Chapter Six. The observed correlation cannot be accounted for by the model which informs the visual-probe task, suggesting that some other significant process is occurring and that the estimates of attentional bias derived from task may be unreliable, particularly when comparing between the two groups. Green circles = patients; blue circles = control group.**

## 7.2. The mechanisms of action of attentional training

Cognitive models of the emotional disorders suggest that abnormalities in the processing of emotional information may cause and maintain the illnesses (Mathews and MacLeod 2005). The most commonly reported abnormality is a tendency for patients with an emotional disorder to preferentially attend to, remember or interpret information in a negative manner; a phenomenon which has been termed "negative emotional bias" (Williams, Watts et al. 1997; Mathews and MacLeod 2005). One of the most convincing demonstrations of the causal role of such biases in the emotional disorders comes from previous work which demonstrated that training healthy individuals to adopt a negative attentional bias increased the depressive and anxious symptoms experienced by those individuals in response to external stress (MacLeod, Rutherford et al. 2002). This seminal work was followed by a series of studies, using a variety of methodologies, which reported that training regimes could also be used to reduce negative attentional biases in clinical populations with primary anxiety disorders (Dandeneau, Baldwin et al. 2007; Hazen, Vasey et al. 2008; Amir, Beard et al. 2009; Amir, Beard et al.

2009; Schmidt, Richey et al. 2009), and non-clinical samples with high rates of depressive symptoms (Wells and Beevers 2009), to beneficial effect. Thus, studies utilising attentional bias training paradigms were able to demonstrate both the causal role of negative attentional bias in the emotional disorders and the possible utility of the training regimes as a treatment in clinical groups with depression and anxiety. The current thesis sought initially to investigate the mechanisms by which this attentional training effect occurred.

### 7.2.1. The level of the attentional training effect

The first question addressed concerned the generalisation of the training effect; in other words, what does attentional training actually train? The answers to this question may be considered along a continuum from specific to general. At the most specific, attentional training may simply influence subsequent attention to the exact stimuli presented in exactly the same manner as the training task. At its most general, attentional bias training may influence the processing of all forms of emotional information, regardless of type or mode of presentation. As the emotional disorders are associated with biased attention, memory and interpretation of a broad range of stimuli type (MacLeod, Mathews et al. 1986; Williams, Watts et al. 1997; Yiend and Mathews 2001; Dandeneau and Baldwin 2004; Gotlib, Krasnoperova et al. 2004; Fox, Russo et al. 2005; Mathews and MacLeod 2005), it seems reasonable to assume that the more general the induced bias, the more likely it is to be clinically relevant. A number of lines of evidence presented in the preceding chapters demonstrate instances in which the training effect was found to generalise. Firstly, the stimuli used in all the assessment tasks consisted of novel exemplars not used in training indicating that the training effect generalises beyond the specific stimuli set used in training (N.B. replicating the finding reported in; MacLeod, Rutherford et al. 2002). Secondly, the effect of

word based training was found to generalise across stimuli type to influence both the behavioural (Chapter Two) and neural (Chapter Three) responses to faces suggesting that the training effect is not limited to the stimuli type used in training. Lastly the effect of repeated, face based training was found to generalise across stimulus duration in Chapter Five and further, to influence alternative cognitive measures such as memory for word stimuli. In all of these examples the direction of effect was congruent with the training received, such that the positive training regimes led to positive biases when compared to negative (Chapters Two and Three) or neutral (Chapter Five) regimes. Overall, these results indicate that training is able to alter the processing of emotional stimuli at a relatively high level, that is training seems to occur to the emotional nature of the stimuli rather than solely to lower level features such as stimuli type.

An alternative approach to assessing the ability of training to alter attention to the emotional dimension of stimuli, not used in the current thesis, would involve modifying the training task such that participants are able to train to a range of different stimuli dimensions. For example, a face based training task could be administered such that the positive face was always female, whereas the negative face was always male. In such a task participants could therefore learn to direct their attention relative to either the emotion, or gender (or both) of the faces. The extent to which each dimension was used to direct attention could then be demonstrated in a follow up task in which the gender and emotion of the faces were no longer confounded. Such a task would provide a within subject measure of the extent to which training occurs to the emotional relative to non-emotional dimensions of stimuli. From a clinical perspective this within subject "emotional preference" measure may provide a useful

tool by which future studies may explore individual differences in the tendency to develop attentional biases (see section 7.2.3 below).

### 7.2.2. Do psychological and pharmacological manipulations of attentional function target distinct control processes?

In Chapter One it was observed that all previous training studies had described the attentional effects (i.e. the outcomes of tasks specifically designed to assess attentional bias) occurring 500 ms or more after stimulus presentation whereas many pharmacological studies had observed attentional effects at shorter durations (including 100 ms; Murphy, Yiend et al. 2009). This was used as evidence to support the proposal that the psychological and pharmacological manipulations of attentional function targeted different processes and, in particular, that antidepressant medication altered the bottom-up, stimulus appraisal system whereas psychological interventions altered the top-down, control systems (Chapter One and; DeRubeis, Siegle et al. 2008; Harmer 2008). This hypothesis was supported by the behavioural findings reported in Chapter Two in which a single episode of training altered the later (500 ms) but not earlier (100 ms) deployment of attention, and by the neuroimaging results from Chapter Three in which a single episode of training was found to alter prefrontal but not amygdala activity. However, the results reported in Chapter Five indicated that the behavioural effects of repeated training did generalise across stimulus duration; in this study training stimuli were presented for relatively long (500 or 1000 ms) durations, but the effect of training was to produce a general alteration in attentional deployment, including a significant effect at a shorter (100 ms) duration. Thus while the influence of a single session of training was limited to longer stimuli durations, repeated training altered attention at earlier durations. This pattern of results is consistent with the initial effect of training being

on top-down control processes but with a transfer to more automatised, bottom-up processes occurring with repeated practice of the training regime. A corollary of this proposal which is amenable to testing in future studies is that the neural effects of a repeated training regime should be detectable across a wider network of regions, including the amygdala.

### 7.2.3. The nature of attentional biases in the emotional disorders

The studies reported in this thesis attest to the flexibility of the attentional control systems. Attentional deployment was found to quickly adapt to both the regularities of training (Chapters Two, Three and Five) and to the recent probabilities with which emotional stimuli were presented (Chapter Four). This reactive flexibility raises interesting questions about the nature of the negative attentional biases which are associated with both depression and anxiety (Williams, Watts et al. 1997; Mathews and MacLeod 2005). For example, why, when faced with the relatively non-threatening stimuli used in previous studies (e.g. unpleasant words), do anxious individuals not adapt as non-anxious individual do by directing attention away from the threatening information? It is not that attentional function in anxiety lacks flexibility as previous training studies have demonstrated this not to be the case (Amir, Beard et al. 2009; Amir, Beard et al. 2009; See, Macleod et al. 2009). Rather, it may be that these individuals are more likely to *react* to certain environments by deploying a negative attentional bias. In this light, the critical measure of attentional function in the emotional disorders may be the readiness with which individuals alter their attentional function in response to environmental information rather than the static, cross-sectional measures of attention deployment generally reported in the literature. Indeed an elegant study by Clarke and colleagues (Clarke, MacLeod et al. 2008) provides experimental support for this proposal by demonstrating that the *change* in attentional function induced by a single episode of

training predicted the increase in anxiety in response to a subsequent naturalistic stressor. The apparent predictive validity of this dynamic measure of attentional function suggests that it may provide an interesting tool for future studies, for example, as an alternative measure of the influence of pharmacological interventions on attention.

*7.2.3.1. Expectation and emotional attention*

The results reported in Chapter Four suggest a novel conceptual framework for approaching the dynamics of attentional deployment to emotional information in anxiety. In Chapter Four the deployment of participants' attention was found to be influenced by a constantly updated assessment of the probability of threat and by the surprise elicited by the violation of this expectation. Applying this insight to the clinical literature suggests two possible routes by which the negative attentional biases found in anxiety (Mathews and MacLeod 2005) may be generated: either patients with anxiety have a greater expectation of threat, or they are more surprised when threatening stimuli appear. Although the former explanation concurs with the clinical presentation of the disorders and is therefore intuitively appealing, the analysis described in Chapter Four suggested that levels of trait anxiety were negatively correlated with utilisation of the threat expectancy signal in the orbitofrontal cortex (OFC) and amygdala and positively correlated with surprise in the left temporal lobe suggesting that anxiety was associated with the inefficient tracking of threat expectancy and thus with greater surprise in response to the stimuli. Although speculative, a possible resolution to this apparent conflict is suggested by previous work indicating the importance of the predictability of the environment to anxiety (Grillon 2002; Mineka and Oehlberg 2008). Specifically, anxiety is reliably generated by exposing organisms to unpredictable environments (Mineka and Oehlberg 2008). Grillon suggests a mechanism for this finding; when threatening stimuli are cued, conditioning occurs to that cue leading to specific fear

learning, whereas in unpredictable environments the lack of predictive cues leads to fear conditioning occurring to the context itself which, in turn, produces anxiety (Grillon 2002). Considering this in terms of the results from Chapter Four described above, anxious individuals may be less able to use the statistical regularities of the environment to predict the occurrence of threatening stimuli. This would lead, in effect, to anxious individuals being chronically exposed to unpredictable environments in which the threatening stimuli are surprising. The increased surprise will lead to persistently increased threat-expectation and, as suggested by Grillon (2002), this will be associated with the general context rather than recent environmental cues. In other words, patients' inability to accurately predict when threatening events are likely to occur in their environment leads them to learn that the environment itself is threatening. As an experimental test of this proposal it would be interesting to assess whether increased use of threat-expectancy to control attention, measured either behaviourally or neurally, reduced the symptoms of anxiety provoked by a subsequent stressor.

## 7.3. Clinical Implications

The second broad goal of the current thesis was to investigate some of the possible clinical implications of attentional bias training.

### 7.3.1. Attentional bias training as a treatment

As reviewed in Chapter One, a number of small scale studies have suggested that positive attentional bias training may reduce the symptoms of anxiety and depression in both clinical and analogue populations (Amir, Beard et al. 2009; Amir, Beard et al. 2009; Schmidt, Richey et al. 2009; See, Macleod et al. 2009; Wells and Beevers 2009). However, not all training regimes are effective (Baert, De Raedt et al. 2010) highlighting the importance of

understanding how it may most effectively be implemented. An advantage of bias modification procedures such as attentional training, over more complex interventions such as CBT, is their suitability for use in controlled experimental studies. This facet of attentional training was utilised in Chapter Five in which the effects of one possible treatment regime, the combination of attentional bias training and antidepressant medication, was assessed. The results of this study did not demonstrate strong support for the particular treatment combination tested, however the study itself provided initial evidence that the use of controlled experimental study designs may be able to complement the standard clinical trial and meta-analytic approaches generally employed to address such questions (Foa, Franklin et al. 2002; Pampallona, Bollini et al. 2004; Cuijpers, van Straten et al. 2009). A second aspect of this study which warrants mention is the demonstration that the training regime itself need not be uniform: In the study participants were presented with a range of emotional stimuli (positive, neutral and negative faces), in a range of combinations (positive/neutral, positive/negative, negative/neutral) for a range of different durations (500 or 1000 ms) whereas the majority of previous training studies have presented a much more restricted range of trials (e.g. neutral and negative words presented for 500 ms; MacLeod, Rutherford et al. 2002). The observation that a training effect was produced in this study, and critically, that this effect generalised over all of the different stimuli and timing parameters indicates that more varied regimes are effective at producing general alterations of attentional function. At a pragmatic level, participants frequently report that attentional training is tedious to complete, indicating that its acceptability as a possible future treatment is likely to be increased if it could be made more engaging. One method of doing this would be to increase the variability of the regime; the current results provide reassurance that it is possible to do so while maintaining the basic training effect.

### 7.3.2. The challenge of translational studies in clinical populations

The second clinically relevant question addressed in the current thesis was whether attentional bias training, which has demonstrable effect on both attentional function and emotional reactivity in individuals with high levels of anxiety or depression (Amir, Beard et al. 2009; Amir, Beard et al. 2009; Schmidt, Richey et al. 2009; See, Macleod et al. 2009; Wells and Beevers 2009), could be usefully deployed in patients with bipolar disorder (Chapter Six). The study demonstrated some of the difficulties in translating experimental techniques, such as the assessment of emotional attention, to patients with serious mental illness. Specifically, the patient group displayed a generally altered performance on the visual-probe task, which was reflected in a greater overall reaction time. As discussed in the chapter, the interpretation of data from this task depends on a number of assumptions and it is not clear that these were being met. Thus, while cognitive models of brain function provide a particularly promising framework for the exploration of aetiological processes in clinical groups, considerable care and some ingenuity is required to ensure that a given task measures the same processes in clinical groups as it does in non-clinical groups.

## 7.4. Limitations

In the penultimate section of this chapter some of the methodological limitations of the thesis are discussed.

### 7.4.1. The assessment of attentional function

The visual-probe task was used as the behavioural measure of attentional deployment in the current thesis. However, as reviewed in Chapter One, there are a large number of other measures of attentional deployment described in the literature. The interpretation of the effects of training on attentional function reported in this thesis may have been strengthened

by the use of some of these complementary measures, such as eye-tracking technology which provides a direct assay of overt attention and has been demonstrated to be sensitive to training interventions (Wadlinger and Isaacowitz 2008). This limitation was particularly apparent in Chapter Five in which participants completed a large number of training sessions, all of which were based on the visual-probe task. In this chapter it seems likely that the sensitivity of the post-training assessment of attentional deployment was reduced by practice effects and that the use of a novel assay of attention would have been particularly useful. Thus future training studies may benefit from using a range of assessments of attentional deployment, ideally including measures which are procedurally distinct from the training task.

### 7.4.2. The use of healthy volunteers in treatment studies

A second concern regarding the studies reported in the current thesis is that they largely involved non-clinical samples, raising the possibility that the results are not applicable to clinical populations. While the central importance of studying the effects of treatment in the appropriate clinical population is beyond contention, healthy volunteer models have been successful in describing the effects of pharmacological (Harmer, Hill et al. 2003; Harmer, Shelley et al. 2004; Browning, Reid et al. 2007) and psychological (MacLeod, Rutherford et al. 2002; Lang, Moulds et al. 2009) interventions and, critically, have produced similar results to those observed in clinical populations (Harmer, O'Sullivan et al. 2009; Steel, Wykes et al. 2010). Thus, assessing the impact of treatments on cognitive measures in non-clinical populations would seem to be an efficient strategy for exploring the mechanisms underlying the treatments. Further, as discussed in Chapter One, an advantage of non-clinical participants is that, in contrast to clinical populations, the treatments tend not to alter baseline

mood or anxiety which may confound the interpretation of the treatment effect. Thus the use of healthy volunteers may provide important complementary data on treatment mechanisms which can be used in the early stages of treatment design.

### 7.4.3. Inferring causality using associational designs

As a central concern of the current thesis has been to assess the mechanisms by which attentional training works, it is important to note the limitations to the inferences which may be drawn from the methodologies used. The majority of the studies in the thesis have used a randomised controlled design with participants being allocated at random to a particular treatment group, thus it is reasonable to infer that the neural and behavioural effects reported arise as a consequence of the specific intervention assigned to their group. However, it is not possible to be certain as to the specific roles of the cognitive systems identified. As an example, Chapter Three described alterations of lateral prefrontal activity as a consequence of training; the design of the study allows confidence that training led to these changes, but it provides only weak evidence that alteration of prefrontal function is causally important to the training effect. These concerns motivated the connectivity analysis reported in the same chapter, and while the results of this analysis support the predictions regarding causality they do not provide unequivocal evidence. Ultimately, the causal role of the lateral prefrontal cortices (or any other neural or cognitive system) in the training effect would most convincingly be demonstrated by showing that the targeted disruption of the function of these areas leads to abolition of the training effect. Interestingly, early data has been published indicating that it may be possible to use transcranial magnetic stimulation (TMS) techniques to alter both lateral prefrontal function and attentional control to threat (Leyman, De Raedt et

al. 2009), suggesting that a similar approach may be feasible in testing the role of these areas in the attentional training effect.

## 7.5. Future Directions

This thesis has investigated the mechanisms underlying some of the treatments of the emotional disorders using an experimental cognitive neuroscience approach. A number of novel lines of inquiry are suggested by the studies reported, many of which may be amenable to testing using similar methodologies. In this final section, some of these possibilities are discussed.

The most straightforward extension to the work described in the thesis would involve characterisation of the neural effects of both single and multisession training regimes in clinical groups, such as patients with anxiety disorders. A related approach would be to assess whether other forms of bias training are mediated by similar mechanisms. For example, is interpretative bias training (Mathews and MacLeod 2002; Holmes, Lang et al. 2009), which encourages individuals to habitually interpret ambiguous information in a particular manner (e.g. the glass is half full rather than half empty), mediated by similar or distinct processes to those recruited in attentional bias training? And, if different, does this translate into a differing pattern of clinical efficacy or interactions with other treatments? The ultimate attraction of both of these lines of enquiry is that a fuller characterisation of the effects of the various cognitive (and pharmacological) interventions available is likely to be an essential step if the neurocognitive model of treatment effect described in this thesis is to be developed to the point that it is able to predict clinically useful outcomes such as treatment efficacy in a clinical group or even, conceivably, an individual.

Lastly, it would be interesting to extend the approach documented in Chapter Five, where the interaction between attentional bias training and antidepressant medication was assessed, to incorporate other determinants of clinical outcome. This may include testing different forms of training or medication, as discussed in Chapter Five, but it could also test the impact of other factors such as genetic risk alleles. For example, a widely studied polymorphism of the promoter region of the serotonin transporter protein has been found to be associated with altered attentional deployment to emotional information (Fox, Ridgewell et al. 2009) as well as increased risk of depression in those who have experienced significant external stress (Caspi, Sugden et al. 2003; although see also, Risch, Herrell et al. 2009). It is possible that attentional bias training could be used to investigate this interaction in an experimental setting. Initially this may involve testing for a genotype x training interaction using similar cognitive and emotional outcomes as described in Chapter Five.

## 7.6. Conclusion

The current thesis has utilised a novel cognitive intervention, attentional bias training, to probe the neural and cognitive systems which control attention to threatening information and to assess novel treatment strategies in both clinical and non-clinical populations. The results indicate that the flexible nature of the attentional control systems may be co-opted by simple computerised tasks and that this initially involves alteration of prefrontal control functions. From a treatment perspective, antidepressant medication was found to interfere with the cognitive effects of attentional training and an initial study found no evidence to support the use of attentional training in bipolar disorder. More broadly the work completed in this thesis demonstrates the utility of using a cognitive-neuroscientific framework to explore the mechanisms of psychiatric treatments and to assess the impact of novel therapeutic strategies.

# References


Adolphs, R. (2010). "What does the amygdala contribute to social cognition?" *Ann N Y Acad Sci* 1191(1): 42-61.

Adolphs, R., Tranel, D., Damasio, H. and Damasio, A. R. (1995). "Fear and the human amygdala." *Journal of Neuroscience* 15(9): 5879-5891.

Algom, D., Chajut, E. and Lev, S. (2004). "A rational look at the emotional Stroop phenomenon: A generic slowdown, not a Stroop effect." *Journal of Experimental Psychology* 133: 323-338.

Almeida, J. R., Versace, A., Hassel, S., Kupfer, D. J. and Phillips, M. L. (2010). "Elevated amygdala activity to sad facial expressions: a state marker of bipolar but not unipolar depression." *Biol Psychiatry* 67(5): 414-21.

Amir, N., Beard, C., Burns, M. and Bomyea, J. (2009). "Attention modification program in individuals with generalized anxiety disorder." *Journal of Abnormal Psychology* 118(1): 28-33.

Amir, N., Beard, C., Taylor, C. T., Klumpp, H., Elias, J., Burns, M. and Chen, X. (2009). "Attention training in individuals with generalized social phobia: A randomized controlled trial." *J Consult Clin Psychol* 77(5): 961-73.

Anderson, A. K. (2005). "Affective influences on the attentional dynamics supporting awareness." *J Exp Psychol Gen* 134(2): 258-81.

Anderson, N. (1968). "Likableness ratings of 555 personality trait words." *J Pers Soc Psychol* 9: 272–279.

APA (1994). *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC, American Psychiatric Association.

Arce, E., Simmons, A. N., Lovero, K. L., Stein, M. B. and Paulus, M. P. (2008). "Escitalopram effects on insula and amygdala BOLD activation during emotional processing." *Psychopharmacology (Berl)* 196(4): 661-672.

Asmundson, G. J. G. and Stein, M. B. (1994). "Selective processing of social threat in patients with generalized social phobia: Evaluation using a dot-probe paradigm." *Journal of Anxiety Disorders* 8(2): 107-117.

Aston-Jones, G. and Cohen, J. D. (2005). "An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance." *Annual Review of Neuroscience* 28: 403-450.

Baert, S., De Raedt, R., Schacht, R. and Koster, E. H. (2010). "Attentional bias training in depression: Therapeutic effects depend on depression severity." *J Behav Ther Exp Psychiatry*.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York.

Beck, A. T., Brown, G. K. and Steer, R. A. (1997). "Psychometric characteristics of the Scale for Suicide Ideation with psychiatric outpatients." *Behaviour, Research and Therapy* 35(11): 1039-1046.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. and Erbaugh, J. (1961). "An inventory for measuring depression." *Archives of General Psychiatry* 4: 53-63.

Becker, E. S., Rinck, M., Margraf, J. and Roth, W. T. (2001). "The emotional Stroop effect in anxiety disorders: general emotional or disorder specificity?" *J Anxiety Disord* 15(3): 147-59.

Behrens, T. E., Hunt, L. T., Woolrich, M. W. and Rushworth, M. F. (2008). "Associative learning of social value." *Nature* 456(7219): 245-9.

Behrens, T. E., Woolrich, M. W., Walton, M. E. and Rushworth, M. F. (2007). "Learning the value of information in an uncertain world." *Nat Neurosci* 10(9): 1214-21.

Bentall, R. P. and Thompson, M. (1990). "Emotional Stroop performance and the manic defence." *British Journal of Clinical Psychology* 29(2): 235-237.

Bishop, S. J. (2007). "Neurocognitive mechanisms of anxiety: an integrative account." *Trends in Cognitive Science* 11(7): 307-316.

Bishop, S. J., Duncan, J., Brett, M. and Lawrence, A. D. (2004a). "Prefrontal cortical function and anxiety: controlling attention to threat-related stimuli." *Nature Neuroscience* 7(2): 184-188.

Bishop, S. J., Duncan, J. and Lawrence, A. D. (2004b). "State Anxiety Modulation of the Amygdala Response to Unattended Threat-Related Stimuli." *Journal of Neuroscience* 24(46): 10364-10368.

Bishop, S. J., Jenkins, R. and Lawrence, A. D. (2007). "Neural processing of fearful faces: effects of anxiety are gated by perceptual capacity limitations." *Cereb Cortex* 17(7): 1595-603.

Blackwell, S. E. and Holmes, E. A. (2010). "Modifying Interpretation and Imagination in Clinical Depression: A Single Case Series using Cognitive Bias Modification." *Applied Cognitive Psychology* in press.

Blier, P. and de Montigny, C. (1994). "Current advances and trends in the treatment of depression." *Trends Pharmacol Sci* 15(7): 220-6.

Booij, L., Van der Does, A. J., Haffmans, P. M., Riedel, W. J., Fekkes, D. and Blom, M. J. (2005). "The effects of high-dose and low-dose tryptophan depletion on mood and

cognitive functions of remitted depressed patients." *Journal of Psychopharmacology* 19(3): 267-275.

Booij, L., Van der Does, A. J. and Riedel, W. J. (2003). "Monoamine depletion in psychiatric and healthy populations: review." *Molecular Psychiatry* 8(12): 951-973.

Botvinick, M. M. (2007). "Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function." *Cogn Affect Behav Neurosci* 7(4): 356-66.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. and Cohen, J. D. (2001). "Conflict monitoring and cognitive control." *Psychological Review* 108: 624-652.

Botvinick, M. M., Cohen, J. D. and Carter, C. S. (2004). "Conflict monitoring and anterior cingulate cortex: an update." *Trends Cogn Sci* 8(12): 539-46.

Boyer, M. C., Compas, B. E., Stanger, C., Colletti, R. B., Konik, B. S., Morrow, S. B. and Thomsen, A. H. (2006). "Attentional biases to pain and social threat in children with recurrent abdominal pain." *J Pediatr Psychol* 31(2): 209-20.

Bradley, B. P., Mogg, K. and 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., Bonham-Carter, C., Fergusson, E., Jenkins, J. and Parr, M. (1997). "Attentional biases for emotional faces." *Cognition and Emotion* 11(1): 25-42.

Brody, A. L., Saxena, S., Stoessel, P., Gillies, L. A., Fairbanks, L. A., Alborzian, S., Phelps, M. E., Huang, S. C., Wu, H. M., Ho, M. L., Ho, M. K., Au, S. C., Maidment, K. and Baxter, L. R., Jr. (2001). "Regional brain metabolic changes in patients with major

depression treated with either paroxetine or interpersonal therapy: preliminary findings." *Arch Gen Psychiatry* 58(7): 631-40.

Brotman, M. A., Rich, B. A., Schmajuk, M., Reising, M., Monk, C. S., Dickstein, D. P., Mogg, K., Bradley, B. P., Pine, D. S. and Leibenluft, E. (2007). "Attention bias to threat faces in children with bipolar disorder and comorbid lifetime anxiety disorders." *Biological Psychiatry* 61(6): 819-821.

Browning, M., Reid, C., Cowen, P. J., Goodwin, G. M. and Harmer, C. J. (2007). "A single dose of citalopram increases fear recognition in healthy subjects." *Journal of Psychopharmacology* 21(7): 684-690.

Bush, G., Luu, P. and Posner, M. I. (2000). "Cognitive and emotional influences in anterior cingulate cortex." *Trends in Cognitive Science* 4(6): 215-222.

Carver, C. S., Johnson, S. L. and 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): 912-943.

Carver, C. S. and White, T. L. (1994). "Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales." *Journal of Personality and Social Psychology* 67: 319-333.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., McClay, J., Mill, J., Martin, J., Braithwaite, A. and Poulton, R. (2003). "Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene." *Science* 301(5631): 386-9.

Chandler, R. A., Wakeley, J., Goodwin, G. M. and Rogers, R. D. (2009). "Altered risk-aversion and risk-seeking behavior in bipolar disorder." *Biol Psychiatry* 66(9): 840-6.

Clark, D. M. (1983). "On the induction of depressed mood in the laboratory: Evaluation and comparison of the velten and musical procedures " *Advances in Behaviour Research and Therapy* 5: 27.

Clark, L., Iversen, S. D. and Goodwin, G. M. (2002). "Sustained attention deficit in bipolar disorder." *Br J Psychiatry* 180: 313-9.

Clark, L. and Sahakian, B. (2006). Neuropsychological and biological approaches to understanding bipolar disorder. *The Psychology of Bipolar Disorder: New Developments and Research Strategies* S. H. Jones and R. P. Bentall. Oxford, Oxford Univeristy Press**:** 139-178.

Clarke, P., MacLeod, C. M. and Shirazee, N. (2008). "Prepared for the worst: Readiness to acquire threat bias and susceptibility to elevate trait anxiety." *Emotion* 8(1): 47-57.

Colom, F., Vieta, E., Martinez-Aran, A., Reinares, M., Goikolea, M., Benabarre, A., Torrent, C., Comes, M., Corbella, B., Parramon, G. and Corominas, J. (2003). " A randomized trial on the efficacy of group psychoeducation in the prophylaxis of recurrences in bipolar patients whose disease is in remission." *Archives of General Psychiatry* 60: 402-407.

Cools, R., Calder, A. J., Lawrence, A. D., Clark, L., Bullmore, E. and Robbins, T. W. (2005). "Individual differences in threat sensitivity predict serotonergic modulation of amygdala response to fearful faces." *Psychopharmacology (Berl)* 180(4): 670-679.

Cooper, R. M. and Langton, S. R. H. (2006). "Attentional bias to angry faces using the dot-probe task? It depends when you look for it." *Behaviour, Research and Therapy* 44(9): 1321-1329.

Corlett, P. R., Aitken, M. R., Dickinson, A., Shanks, D. R., Honey, G. D., Honey, R. A., Robbins, T. W., Bullmore, E. T. and Fletcher, P. C. (2004). "Prediction error during

retrospective revaluation of causal associations in humans: fMRI evidence in favor of an associative model of learning." *Neuron* 44(5): 877-88.

Cuijpers, P., van Straten, A., Warmerdam, L. and Andersson, G. (2009). "Psychotherapy versus the combination of psychotherapy and pharmacotherapy in the treatment of depression: a meta-analysis." *Depress Anxiety* 26(3): 279-88.

Dandeneau, S. D. and Baldwin, M. W. (2004). "The Inhibition of Socially Rejecting Information Among People with High Versus Low Self-Esteem: The Role of Attentional Bias and the Effects of Bias Reduction Training." *Journal of Social and Clinical Psychology* 23(4): 584-602.

Dandeneau, S. D., Baldwin, M. W., Baccus, J. R., Sakellaropoulo, M. and Pruessner, J. C. (2007). "Cutting stress off at the pass: reducing vigilance and responsiveness to social threat by manipulating attention." *Journal of Personality and Social Psychology* 93(4): 651-666.

Dannlowski, U., Kersting, A., Lalee-Mentzel, J., Donges, U. S., Arolt, V. and Suslow, T. (2006). "Subliminal affective priming in clinical depression and comorbid anxiety: a longitudinal investigation." *Psychiatry Res* 143(1): 63-75.

Davis, M. and Whalen, P. J. (2001). "The amygdala: vigilance and emotion." *Molecular Psychiatry* 6(1): 13-34.

De Martino, B., Strange, B. A. and Dolan, R. J. (2008). "Noradrenergic neuromodulation of human attention for emotional and neutral stimuli." *Psychopharmacology (Berl).* 197(1): 127-136.

De Raedt, R. and Koster, E. H. (2010). "Understanding vulnerability for depression from a cognitive neuroscience perspective: A reappraisal of attentional factors and a new conceptual framework." *Cogn Affect Behav Neurosci* 10(1): 50-70.

Deichmann, R., Gottfried, J. A., Hutton, C. and Turner, R. (2003). "Optimized EPI for fMRI studies of the orbitofrontal cortex." *Neuroimage* 19(2 Pt 1): 430-441.

Del-Ben, C. M., Deakin, J. F., McKie, S., Delvai, N. A., Williams, S. R., Elliott, R., Dolan, M. and Anderson, I. M. (2005). "The effect of citalopram pretreatment on neuronal responses to neuropsychological tasks in normal volunteers: an FMRI study." *Neuropsychopharmacology* 30(9): 1724-1734.

den Ouden, H. E., Friston, K. J., Daw, N. D., McIntosh, A. R. and Stephan, K. E. (2009). "A dual role for prediction error in associative learning." *Cereb Cortex* 19(5): 1175-85.

Derakshan, N., Salt, M. and Koster, E. H. (2009). "Attentional control in dysphoria: an investigation using the antisaccade task." *Biol Psychol* 80(2): 251-5.

Derryberry, D. and Reed, M. A. (2002). "Anxiety-related attentional biases and their regulation by attentional control." *Journal of Abnormal Psychology* 111(2): 225-36.

DeRubeis, R. J., Hollon, S. D., Amsterdam, J. D., Shelton, R. C., Young, P. R., Salomon, R. M., O'Reardon, J. P., Lovett, M. L., Gladis, M. M., Brown, L. L. and Gallop, R. (2005). "Cognitive therapy vs medications in the treatment of moderate to severe depression." *Archives of General Psychiatry* 62(4): 409-416.

DeRubeis, R. J., Siegle, G. J. and Hollon, S. D. (2008). "Cognitive therapy versus medication for depression: treatment outcomes and neural mechanisms." *Nature Reviews Neuroscience* 9(10): 788-796.

Desimone, R. and Duncan, J. (1995). "Neural Mechanisms of Selective Visual Attention." *Annual Review of Neuroscience* 18: 193-222.

Donaldson, C., Lam, D. and Mathews, A. (2007). "Rumination and attention in major depression." *Behaviour, Research and Therapy* 45(11): 2664-2678.

Dozois, D. J. and Dobson, K. S. (2001). "Information processing and cognitive organization in unipolar depression: specificity and comorbidity issues." *J Abnorm Psychol* 110(2): 236-46.

Driver, J. (2001). "A selective review of selective attention research from the past century." *Br J Psychol* 92(Pt 1): 53-78.

Dudley, R., O'Brien, J., Barnett, N., McGuckin, L. and Britton, P. (2002). "Distinguishing depression from dementia in later life: a pilot study employing the Emotional Stroop task." *Int J Geriatr Psychiatry* 17(1): 48-53.

Duman, R. S. (2004). "Depression: a case of neuronal life and death?" *Biol Psychiatry* 56(3): 140-5.

Duncan, J. and Owen, A. M. (2000). "Common regions of the human frontal lobe recruited by diverse cognitive demands." *Trends Neurosci* 23(10): 475-83.

Egner, T. (2007). "Congruency sequence effects and cognitive control." *Cogn Affect Behav Neurosci* 7(4): 380-90.

Eimer, M., Holmes, A. and McGlone, F. P. (2003). "The role of spatial attention in the processing of facial expression: an ERP study of rapid brain responses to six basic emotions." *Cogn Affect Behav Neurosci* 3(2): 97-110.

Ekman, P. and Friesen, W. V. (1976). *Pictures of Facial Affect*. Palo Alto, Calif, Consulting Psychologists Press.

Eldara, S., Ricona, T. and Bar-Haim, Y. (2008). "Plasticity in attention: Implications for stress response in children." *Behaviour, Research and Therapy* 46(4): 450-461.

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R. and Hirsch, J. (2006). "Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala." *Neuron* 51: 871-882.

Evers, E. A., van derVeen, F. M., Jolles, J., Deutz, N. E. and Schmitt, J. A. (2006). "Acute tryptophan depletion improves performance and modulates the BOLD response during a Stroop task in healthy females." *Neuroimage* 32: 248-255.

Eysenck, M. W., Derakshan, N., Santos, R. and Calvo, M. G. (2007). "Anxiety and cognitive performance: attentional control theory." *Emotion* 7(2): 336-53.

Fales, C. L., Barch, D. M., Rundle, M. M., Mintun, M. A., Mathews, J., Snyder, A. Z. and Sheline, Y. I. (2009). "Antidepressant treatment normalizes hypoactivity in dorsolateral prefrontal cortex during emotional interference processing in major depression." *Journal of Affective Disorders* 112(1-3): 206-211.

Fales, C. L., Barch, D. M., Rundle, M. M., Mintun, M. A., Snyder, A. Z., Cohen, J. D., Mathews, J. and Sheline, Y. I. (2008). "Altered emotional interference processing in affective and cognitive-control brain circuitry in major depression." *Biological Psychiatry* 63(4): 377-384.

Foa, E. B., Franklin, M. E. and Moser, J. (2002). "Context in the clinic: how well do cognitive-behavioral therapies and medications work in combination?" *Biological Psychiatry* 52(10): 987-997.

Fox, E., Ridgewell, A. and Ashwin, C. (2009). "Looking on the bright side: biased attention and the human serotonin transporter gene." *Proceedings of the Royal Society B.* 276(1663): 1747-51.

Fox, E., Russo, R., Bowles, R. and Dutton, K. (2001). "Do threatening stimuli draw or hold visual attention in subclinical anxiety?" *J Exp Psychol Gen.* 130(4): 681-700.

Fox, E., Russo, R. and Georgiou, G. A. (2005). "Anxiety modulates the degree of attentive resources required to process emotional faces." *Cognitive, Affective and Behavioural Neuroscience* 5(4): 396-404.

French, C. C., Richards, A. and Scholfield, E. J. (1996). "Hypomania, anxiety and the emotional Stroop." *British Journal of Clinical Psychology* 35(4): 617-626.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E. and Dolan, R. J. (1997). "Psychophysiological and modulatory interactions in neuroimaging." *Neuroimage* 6(3): 218-29.

Friston, K. J., Kilner, J. and Harrison, L. (2006). "A free energy principle for the brain." *J Physiol Paris* 100(1-3): 70-87.

Fu, C. H., Williams, S. C., Cleare, A. J., Brammer, M. J., Walsh, N. D., Kim, J., Andrew, C. M., Pich, E. M., Williams, P. M., Reed, L. J., Mitterschiffthaler, M. T., Suckling, J. and Bullmore, E. T. (2004). "Attenuation of the neural response to sad faces in major depression by antidepressant treatment: a prospective, event-related functional magnetic resonance imaging study." *Archives of General Psychiatry* 61(9): 877-889.

Fu, C. H., Williams, S. C., Cleare, A. J., Scott, J., Mitterschiffthaler, M. T., Walsh, N. D., Donaldson, C., Suckling, J., Andrew, C., Steiner, H. and Murray, R. M. (2008). "Neural responses to sad facial expressions in major depression following cognitive behavioral therapy." *Biological Psychiatry* 64(6): 505-512.

Furukawa, T. A., Watanabe, N. and Churchill, R. (2007). "Combined psychotherapy plus antidepressants for panic disorder with or without agoraphobia." *Cochrane Database of Systematic Reviews*(1).

Gilboa-Schechtman, E., Erhard-Weiss, D. and Jeczemien, P. (2002). "Interpersonal deficits meet cognitive biases: memory for facial expressions in depressed and anxious men and women." *Psychiatry Research* 113(3): 279-293.

Glascher, J., Hampton, A. N. and O'Doherty, J. P. (2009). "Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making." *Cereb Cortex* 19(2): 483-95.

Goldapple, K., Segal, Z., Garson, C., Lau, M., Bieling, P., Kennedy, S. and Mayberg, H. (2004). "Modulation of cortical-limbic pathways in major depression: treatment-specific effects of cognitive behavior therapy." *Arch Gen Psychiatry* 61(1): 34-41.

Gotlib, I. H., Kasch, K. L., Traill, S. K., Joormann, J., Arnow, B. A. and Johnson, S. L. (2004). "Coherence and specificity of information processing biases in depression and social phobia." *Journal of Abnormal Psychology* 113(386-398).

Gotlib, I. H., Krasnoperova, E., Yue, D. N. and Joormann, J. (2004). "Attentional biases for negative interpersonal stimuli in clinical depression." *Journal of Abnormal Psychology* 113: 127-135.

Grant, D. M. and Beck, J. G. (2006). "Attentional biases in social anxiety and dysphoria: does comorbidity make a difference?" *J Anxiety Disord* 20(4): 520-9.

Grillon, C. (2002). "Startle reactivity and anxiety disorders: aversive conditioning, context, and neurobiology." *Biol Psychiatry* 52(10): 958-75.

Grillon, C. and Baas, J. (2003). "A review of the modulation of the startle reflex by affective states and its application in psychiatry." *Clinical Neurophysiology* 114: 1557-1579.

Hamilton, M. (1960). "A rating scale for depression." *Journal of Neurology, Neurosurgery and Psychiatry* 23: 56-62.

Hampton, A. N., Bossaerts, P. and O'Doherty, J. P. (2008). "Neural correlates of mentalizing-related computations during strategic interactions in humans." *Proc Natl Acad Sci U S A* 105(18): 6741-6.

Hansen, C. H. and Hansen, R. D. (1988). "Finding the face in the crowd: an anger superiority effect." *J Pers Soc Psychol* 54(6): 917-24.

Harmer, C. J. (2008). "Serotonin and emotional processing: Does it help explain antidepressant drug action?" *Neuropharmacology* 55(6): 1023-1028.

Harmer, C. J., Bhagwagar, Z., Perrett, D. I., Vollm, B. A., Cowen, P. J. and Goodwin, G. M. (2003). "Acute SSRI administration affects the processing of social cues in healthy volunteers." *Neuropsychopharmacology* 28(1): 148-152.

Harmer, C. J., Hill, S. A., Taylor, M. J., Cowen, P. J. and Goodwin, G. M. (2003). "Toward a neuropsychological theory of antidepressant drug action: increase in positive emotional bias after potentiation of norepinephrine activity." *Am J Psychiatry* 160(5): 990-992.

Harmer, C. J., Mackay, C. E., Reid, C. B., Cowen, P. J. and Goodwin, G. M. (2006). "Antidepressant drug treatment modifies the neural processing of nonconscious threat cues." *Biological Psychiatry* 59(9): 816-20.

Harmer, C. J., O'Sullivan, U., Favaron, E., Massey-Chase, R., Ayres, R., Reinecke, A., Goodwin, G. M. and Cowen, P. J. (2009). "Effect of acute antidepressant administration on negative affective bias in depressed patients." *Am J Psychiatry* 166(10): 1178-84.

Harmer, C. J., Shelley, N. C., Cowen, P. J. and Goodwin, G. M. (2004). "Increased positive versus negative affective perception and memory in healthy volunteers following selective serotonin and norepinephrine reuptake inhibition." *Am J Psychiatry* 161(7): 1256-1263.

Harvey, A., Watkins, E. R., Mansell, W. and Shafran, R. (2004). *Cognitive Behavioural Processes across Psychological Disorders: A transdiagnostic approach to research and treatment*. Oxford, Oxford University Press.

Hayes, S., Hirsch, C. R. and Mathews, A. (2010). "Facilitating a benign attentional bias reduces negative thought intrusions." *J Abnorm Psychol* 119(1): 235-40.

Hayward, G., Goodwin, G. M., Cowen, P. J. and Harmer, C. J. (2005). "Low-dose tryptophan depletion in recovered depressed patients induces changes in cognitive processing without depressive symptoms." *Biological Psychiatry* 57(5): 517-524.

Hazen, R. A., Vasey, M. W. and Schmidt, N. B. (2008). "Attentional retraining: A randomized clinical trial for pathological worry." *Journal of Psychiatric Research* 43(6): 627-633.

Hertel, P. and Brozovich, F. (2010). "Cognitive Habits and Memory Distortions in Anxiety and Depression." *Current Directions in Psycholoigcal Science*.

Hogg, R. V., McKean, J. W. and Craig, A. T. (2005). *Introduction to Mathematical Statistics (International Edition)*, Pearson Education.

Hollon, S. D., Stewart, M. O. and Strunk, D. (2006). "Enduring effects for cognitive behavior therapy in the treatment of depression and anxiety." *Annual Review of Psychology* 57: 285-315.

Holmes, E. A., Lang, T. J. and Shah, D. M. (2009). "Developing interpretation bias modification as a "cognitive vaccine" for depressed mood: imagining positive events makes you feel better than thinking about them verbally." *J Abnorm Psychol* 118(1): 76-88.

Hope, D. A., Rapee, R. M., Heimberg, R. G. and Dombeck , M. J. (1990). "Representations of the self in social phobia: Vulnerability to social threat." *Cognitive Therapy and Research* 14(2): 177-189.

Hopfinger, J. B., Buonocore, M. H. and Mangun, G. R. (2000). "The neural mechanisms of top-down attentional control." *Nat Neurosci* 3(3): 284-91.

Itti, L. and Baldi, P. (2009). "Bayesian surprise attracts human attention." *Vision Res* 49(10): 1295-306.

James, W. (1890). *The Principles of Psychology*. New York, Holt.

Jones, A. D., Cho, R. Y., Nystrom, L. E., Cohen, J. D. and Braver, T. S. (2002). "A computational model of anterior cingulate function in speeded response tasks: effects of frequency, sequence, and conflict." *Cogn Affect Behav Neurosci* 2(4): 300-17.

Jones, L., Scott, J., Haque, S., Gordon-Smith, K., Heron, J., Caesar, S., Cooper, C., Forty, L., Hyde, S., Lyon, L., Greening, J., Sham, P., Farmer, A., McGuffin, P., Jones, I. and Craddock, N. (2005). "Cognitive style in bipolar disorder." *Br J Psychiatry* 187: 431-7.

Jongen, E. M., Smulders, F. T., Ranson, S. M., Arts, B. M. and Krabbendam, L. (2007). "Attentional bias and general orienting processes in bipolar disorder." *Journal of Behavioural Therapy and Experimental Psychiatry* 38(2): 168-183.

Joormann, J. and Gotlib, I. H. (2007). "Selective attention to emotional faces following recovery from depression." *Journal of Abnormal Psychology* 116(1): 80-85.

Joormann, J., Talbot, L. and Gotlib, I. H. (2007). "Biased processing of emotional information in girls at risk for depression." *J Abnorm Psychol* 116(1): 135-43.

Kanwisher, N., McDermott, J. and Chun, M. M. (1997). "The fusiform face area: a module in human extrastriate cortex specialized for face perception." *Journal of Neuroscience* 17(11): 4302-4311.

Kauer-Sant'Anna, M., Kapczinski, F. and Vieta, E. (2009). "Epidemiology and management of anxiety in patients with bipolar disorder." *CNS Drugs* 23(11): 953-64.

Keck, P. E., Strawn, J. R. and McElroy, S. L. (2006). "Pharmacologic treatment considerations in co-occurring bipolar and anxiety disorders." *Journal of Clinical Psychiatry* 67(Suppl 1): 8-15.

Kennedy, S. H., Evans, K. R., Krüger, S., Mayberg, H. S., Meyer, J. H., McCann, S., Arifuzzman, A. I., Houle, S. and Vaccarino, F. J. (2001). "Changes in regional brain glucose metabolism measured with positron emission tomography after paroxetine treatment of major depression." *American Journal of Psychiatry* 158(6): 899-905.

Kerestes, R., Labuschagne, I., Croft, R. J., O'Neill, B. V., Bhagwagar, Z., Phan, K. L. and Nathan, P. J. (2009). "Evidence for modulation of facial emotional processing bias during emotional expression decoding by serotonergic and noradrenergic antidepressants: an event-related potential (ERP) study." *Psychopharmacology* 202(4): 621-34.

Kerr, N., Scott, J. and Phillips, M. L. (2005). "Patterns of attentional deficits and emotional bias in bipolar and major depressive disorder." *British Journal of Clinical Psychology* 44(3): 343-356.

Kessler, R. (1999). Comorbidity of unipolar and bipolar depression with other psychiatric disorders in a general population survey. *Comorbidity in Affective Disorders*. M. Tohen. New York, Marcel Dekker Inc.

Koster, E. H., Crombez, G., Verschuere, B., Van Damme, S. and Wiersema, J. R. (2006). "Components of attentional bias to threat in high trait anxiety: Facilitated engagement, impaired disengagement, and attentional avoidance." *Behav Res Ther* 44(12): 1757-71.

Koster, E. H., Crombez, G., Verschuere, B., Vanvolsem, P. and De Houwer, J. (2007). "A time-course analysis of attentional cueing by threatening scenes." *Exp Psychol* 54(2): 161-71.

Krebs, G., Hirsch, C. R. and Mathews, A. (2010). "The effect of attention modification with explicit vs. minimal instructions on worry." *Behav Res Ther* 48(3): 251-6.

Kringelbach, M. L. (2005). "The human orbitofrontal cortex: linking reward to hedonic experience." *Nat Rev Neurosci* 6(9): 691-702.

Kringelbach, M. L. and Rolls, E. T. (2004). "The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology." *Prog Neurobiol* 72(5): 341-72.

Kushner, M. G., Kim, S. W., Donahue, C., Thuras, P., Adson, D., Kotlyar, M., McCabe, J., Peterson, J. and Foa, E. B. (2007). "D-cycloserine augmented exposure therapy for obsessive-compulsive disorder." *Biological Psychiatry* 62(8): 835-8.

Lam, D. H., Hayward, P., Watkins, E. R., Wright, K. and Sham, P. (2005). "Relapse Prevention in Patients With Bipolar Disorder: Cognitive Therapy Outcome After 2 Years." *American Journal of Psychiatry* 162: 324-329.

Lang, T. J., Moulds, M. L. and Holmes, E. A. (2009). "Reducing depressive intrusions via a computerized cognitive bias modification of appraisals task: developing a cognitive vaccine." *Behav Res Ther* 47(2): 139-45.

Larsen, R. J., Mercer, K. A. and Balota, D. A. (2006). "Lexical characteristics of words used in emotional Stroop experiments." *Emotion* 6(1): 62-72.

Lavie, N. (2005). "Distracted and confused?: selective attention under load." *Trends Cogn Sci* 9(2): 75-82.

Lee, M. and Shafran, R. (2004). "Information processing biases in eating disorders." *Clin Psychol Rev* 24(2): 215-38.

Leyman, L., De Raedt, R., Schacht, R. and Koster, E. H. (2007). "Attentional biases for angry faces in unipolar depression." *Psychological Medicine* 37(3): 393-402.

Leyman, L., De Raedt, R., Vanderhasselt, M. A. and Baeken, C. (2009). "Influence of high-frequency repetitive transcranial magnetic stimulation over the dorsolateral prefrontal cortex on the inhibition of emotional information in healthy volunteers." *Psychol Med* 39(6): 1019-28.

Li, S., Tan, J., Qian, M. and Liu, X. (2008). "Continual training of attentional bias in social anxiety." *Behav Res Ther* 46(8): 905-12.

Lim, S. L. and Kim, J. H. (2005). "Cognitive processing of emotional information in depression, panic, and somatoform disorder." *J Abnorm Psychol* 114(1): 50-61.

Lim, S. L., Padmala, S. and Pessoa, L. (2009). "Segregating the significant from the mundane on a moment-to-moment basis via direct and indirect amygdala contributions." *Proc Natl Acad Sci U S A* 106(39): 16841-6.

Logan, G. D. and Zbrodoff, N. J. (1979). "When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task." *Memory and Cognition* 7: 166-174.

Luck, S. J., Chelazzi, L., Hillyard, S. A. and Desimone, R. (1997). "Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex." *J Neurophysiol* 77(1): 24-42.

Lundqvist, D., Flykt, A. and Ohman, A. (1998). KDEF. Stockholm, Karolinska Hospital.

Lynch, D., Laws, K. R. and McKenna, P. J. (2010). "Cognitive behavioural therapy for major psychiatric disorder: does it really work? A meta-analytical review of well-controlled trials." *Psychol Med* 40(1): 9-24.

Lyon, H. M., Startup, M. and Bentall, R. P. (1999). "Social cognition and the manic defense: attributions, selective attention, and self-schema in bipolar affective disorder." *Journal of Abnormal Psychology* 108(2): 273-282.

MacDonald, A. W., Cohen, J. D., Stenger, V. A. and Carter, C. S. (2000). "Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control." *Science* 288(5472): 1835-1838.

MacKay, D. J. C. (2003). *Information Theory, Inference and Learning Algorithms*. Cambridge, Cambridge University Press.

MacLeod, C. M., Campbell, L. W., Rutherford, E. and Wilson, E. (2004). The causal status of anxiety linked attentional and interpretive bias. *Cognition, Emotion and Psychopathology*. J. Yiend. Cambridge, Cambridge University Press.

MacLeod, C. M. and Hagan, R. (1992). "Individual differences in the selective processing of threatening information, and emotional responses to a stressful life event." *Behaviour, Research and Therapy* 30(2): 151-161.

MacLeod, C. M., Koster, E. H. and Fox, E. (2009). "Whither cognitive bias modification research? Commentary on the special section articles." *Journal of Abnormal Psychology* 118(1): 89-99.

MacLeod, C. M. and Mathews, A. (1988). "Anxiety and the allocation of attention to threat." *Quart. J. Exp. Psychol: Hum. Exp. Psychol* 38: 659-610.

MacLeod, C. M., Mathews, A. and Tata, P. (1986). "Attentional bias in emotional disorders." *Journal of Abnormal Psychology* 95: 15-20.

MacLeod, C. M., Rutherford, E., Campbell, L., Ebsworthy, G. and 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): 107-123.

Martin, S. D., Martin, E., Rai, S. S., Richardson, M. A. and Royall, R. (2001). "Brain blood flow changes in depressed patients treated with interpersonal psychotherapy or venlafaxine hydrochloride: preliminary findings." *Arch Gen Psychiatry* 58(7): 641-8.

Maser, J. D. and Cloninger, C. R. (1990). *Comorbidity of Mood and Anxiety Disorders*. Washington, D.C., American Psychiatric Press.

Mathews, A. and Mackintosh, B. (1998). "A cognitive model of selective processing in anxiety." *Cognitive Therapy and Research* 22(6): 539-560.

Mathews, A. and MacLeod, C. (1986). "Discrimination of threat cues without awareness in anxiety states." *Journal of Abnormal Psychology* 95(2): 131-138.

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

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

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

Mathews, A., Mogg, K., Kentish, J. and Eysenck, M. (1995). "Effect of psychological treatment on cognitive bias in generalized anxiety disorder." *Behaviour, Research and Therapy* 33(3): 293-303.

Mathews, A., Mogg, K., May, J. and Eysenck, M. (1989). "Implicit and explicit memory bias in anxiety." *J Abnorm Psychol* 98(3): 236-40.

Matsumoto, D. and Ekman, P. (1988). Japanese and Caucasian Facial Expressions of Emotion and Neutral Faces (JACFEE and JACNeuF). San Francisco, Available from Human Interaction Laboratory, University of California.

Mattia, J. I., Heimberg, R. G. and Hope, D. A. (1993). "The revised Stroop color-naming task in social phobics." *Behav Res Ther* 31(3): 305-13.

Mayberg, H. S., Liotti, M., Brannan, S. K., McGinnis, S., Mahurin, R. K., Jerabek, P. A., Silva, J. A., Tekell, J. L., Martin, C. C., Lancaster, J. L. and Fox, P. T. (1999). "Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness." *American Journal of Psychiatry* 156(5): 675-682.

McFarlin, D. B. and Blascovich, J. (1984). "On the Remote Associates Test (RAT) as an alternative to illusory performance feedback: A methodological note." *Basic and Applied Social Psychology* 5: 223-229.

McIntyre, R. S. and Keck, P. E. (2006). "Comorbidity in bipolar disorder: clinical and research opportunities." *Bipolar Disorders* 8: 645-647.

McIntyre, R. S., Soczynska, J. K., Bottas, A., Bordbar, K., Konarski, J. Z. and Kennedy, S. H. (2006). "Anxiety disorders and bipolar disorder: a review." *Bipolar Disorders* 8: 665-676.

McRae, K., Hughes, B., Chopra, S., Gabrieli, J. D., Gross, J. J. and Ochsner, K. N. (2009). "The Neural Bases of Distraction and Reappraisal." *Journal of Cognitive Neuroscience*.

Merens, W., Booij, L., Haffmans, P. J. and Van der Does, A. J. W. (2008). "The effects of experimentally lowered serotonin function on emotional information processing and memory in remitted depressed patients." *Journal of Psychopharmacology* 22(6): 653-662.

Miklowitz, D. J., George, E. L., Richards, J. A., Simoneau, T. L. and Suddath, R. L. (2003). "A Randomized Study of Family-Focused Psychoeducation and Pharmacotherapy in the Outpatient Management of Bipolar Disorder." *Archives of General Psychiatry* 60: 904-912.

Mineka, S. and Oehlberg, K. (2008). "The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders." *Acta Psychol (Amst)* 127(3): 567-80.

Mineka, S., Watson, D. and Clark, L. A. (1998). "Comorbidity of anxiety and unipolar mood disorders." *Annual Review of Psychology* 49(377-412).

Moberly, N. J. and Watkins, E. R. (2006). "Processing mode influences the relationship between trait rumination and emotional vulnerability." *Behav Ther* 37(3): 281-91.

Mogg, K. and Bradley, B. P. (1998). "A cognitive-motivational analysis of anxiety." *Behaviour, Research and Therapy* 36(9): 809-848.

Mogg, K., Bradley, B. P. and Williams, R. (1995). "Attentional bias in anxiety and depression: The role of awareness." *British Journal of Clinical Psychology* 34: 17-36.

Monsell, S. (2003). "Task switching." *Trends in Cognitive Science* 7(3): 134-140.

Munafo, M. R., Hayward, G. and Harmer, C. J. (2006). "Selective processing of social threat cues following acute tryptophan depletion." *J Psychopharmacol* 20(1): 33-9.

Munafò, M. R., Hayward, G. and Harmer, C. J. (2006). "Selective processing of social threat cues following acute tryptophan depletion." *Journal of Psychopharmacology* 20(1): 33-39.

Murphy, S. E. (2007). Antidepressants and the early processing of threat. *Department of Psychiatry*, University of Oxford. D.Phil.

Murphy, S. E., Downham, C., Cowen, P. J. and Harmer, C. J. (2008). "Direct effects of diazepam on emotional processing in healthy volunteers." *Psychopharmacology (Berl)* 199(4): 503-513.

Murphy, S. E., Longhitano, C. L., Ayres, R. E., Cowen, P. J. and Harmer, C. J. (2006). "Tryptophan supplementation induces a positive bias in the processing of emotional material in healthy female volunteers." *Psychopharmacology* 187(1): 121-130.

Murphy, S. E., Norbury, R., O'Sullivan, U., Cowen, P. J. and Harmer, C. J. (2009). "Effect of a single dose of citalopram on amygdala response to emotional faces." *British Journal of Psychiatry* 194(6): 535-540.

Murphy, S. E., Yiend, J., Lester, K. J., Cowen, P. J. and Harmer, C. J. (2009). "Short-term serotonergic but not noradrenergic antidepressant administration reduces attentional vigilance to threat in healthy volunteers." *International Journal of Neuropsychopharmacology* 12(2): 169-179.

Murray, E. A. (2007). "The amygdala, reward and emotion." *Trends Cogn Sci* 11(11): 489-97.

Nachev, P., Rees, G., Parton, A., Kennard, C. and Husain, M. (2005). "Volition and conflict in human medial frontal cortex." *Curr Biol* 15(2): 122-8.

Najmi, S. and Amir, N. (2010). "The effect of attention training on a behavioral test of contamination fears in individuals with subclinical obsessive-compulsive symptoms." *J Abnorm Psychol* 119(1): 136-42.

NICE (2006). The management of bipolar disorder in adults, children and adolescents, in primary and secondary care. London, NICE.

NICE (2007). Management of anxiety (panic disorder, with or without agoraphobia, and generalised anxiety disorder) in adults in primary, secondary and community care (amended). London, NICE.

NICE (2009). Treatment and management of depression in adults, including adults with a chronic physical health problem. London, NICE.

Nisbett, R. E. and Wilson, T. D. C. (1977). "Telling more than we can know: Verbal reports on mental processes." *Psychological Review* 84(3): 231-259.

Norberg, M. M., Krystal, J. H. and Tolin, D. F. (2008). "A meta-analysis of D-cycloserine and the facilitation of fear extinction and exposure therapy." *Biological Psychiatry* 63(12): 1118-26.

Norbury, R., Mackay, C. E., Cowen, P. J., Goodwin, G. M. and Harmer, C. J. (2007). "Short-term antidepressant treatment and facial processing. Functional magnetic resonance imaging study." *British Journal of Psychiatry* 190: 531-532.

Nutt, D. J. (2002). "The neuropharmacology of serotonin and noradrenaline in depression." *International Clinical Psychopharmacology* 17(Suppl. 1): S1-S12.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K. J. and Dolan, R. J. (2004). "Dissociable roles of ventral and dorsal striatum in instrumental conditioning." *Science* 304(5669): 452-4.

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J. and Andrews, C. (2001). "Abstract reward and punishment representations in the human orbitofrontal cortex." *Nat Neurosci* 4(1): 95-102.

Ochsner, K. N. and Gross, J. J. (2005). "The cognitive control of emotion." *Trends in Cognitive Science* 9(5): 242-249.

Ochsner, K. N., Ray, R. D., Cooper, J. C., Robertson, E. R., Chopra, S., Gabrieli, J. D. and Gross, J. J. (2004). "For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion." *Neuroimage* 23(2): 483-499.

Öhman, A. (2002). "Automaticity and the amygdala: nonconscoius responses to emotional faces." *Curr. Directions Psycholo. Sci* 11: 62-66.

Ostlund, S. B. and Balleine, B. W. (2007). "Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning." *J Neurosci* 27(18): 4819-25.

Otto, M. W., Simon, N. M., Wisniewski, S. R., Miklowitz, D. J., Kogan, J. N., Reilly-Harrington, N. A., Frank, E., Nierenberg, A. A., Marangell, L. B., Sagduyu, K., Weiss, R. D., Miyahara, S., Thas, M. E., Sachs, G. S., Pollack, M. H. and Investigators., S.-B. (2006). "Prospective 12-month course of bipolar disorder in out-patients with and without comorbid anxiety disorders." *British Journal of Psychiatry* 189: 20-25.

Pampallona, S., Bollini, P., Tibaldi, G., Kupelnick, B. and Munizza, C. (2004). "Combined pharmacotherapy and psychological treatment for depression: a systematic review." *Archives of General Psychiatry* 61(7): 714-9.

Paquette, V., Levesque, J., Mensour, B., Leroux, J. M., Beaudoin, G., Bourgouin, P. and Beauregard, M. (2003). ""Change the mind and you change the brain": effects of

cognitive-behavioral therapy on the neural correlates of spider phobia." *Neuroimage* 18(2): 401-9.

Paulus, M. P., Feinstein, J. S., Castillo, G., Simmons, A. N. and Stein, M. B. (2005). "Dose-dependent decrease of activation in bilateral amygdala and insula by lorazepam during emotion processing." *Archives of General Psychiatry* 62(3): 282-8.

Pessoa, L., Padmala, S. and Morland, T. (2005). "Fate of unattended fearful faces in the amygdala is determined by both attentional resources and cognitive modulation." *NeuroImage* 28: 249-255.

Pessoa, L. and Ungerleider, L. G. (2005). Visual Attention and Emotional Percpetion. *Neurobiology of Attention*. L. Itti, G. Rees and J. K. Tsotsos. Burlington, Elsevier**:** 160-166.

Phan, K. L., Fitzgerald, D. A., Nathan, P. J., Moore, G. J., Uhde, T. W. and Tancer, M. E. (2005). "Neural substrates for voluntary suppression of negative affect: a functional magnetic resonance imaging study." *Biological Psychiatry* 57(3): 210-9.

Phelps, E. A. and Le Doux, J. E. (2005). "Contributions of the Amygdala to Emotion Processing: From Animal Models to Human Behavior." *Neuron* 48: 175-187.

Polivy, J. and Doyle, C. (1980). "Laboratory induction of mood states through the reading of self-referent mood statements: affective changes or demand characteristics?" *J Abnorm Psychol* 89(2): 286-90.

Pope, H. G. J., Cohane, G. H., Kanayama, G., Siegel, A. J. and Hudson, J. I. (2003). "Testosterone gel supplementation formenwith refractory depression: A randomized, placebo-controlled trial." *American Journal of Psychiatry* 160: 105-111.

Posner, M. I. (1980). "Orienting of attention." *Q J Exp Psychol* 32(1): 3-25.

Posner, M. I. and Petersen, S. E. (1990). "The attention system of the human brain." *Annual Review of Neuroscience* 13: 25-42.

Putman, P., Hermans, E. J., Koppeschaar, H., van Schijndel, A. and van Honk, J. (2007). "A single administration of cortisol acutely reduces preconscious attention for fear in anxious young men." *Psychoneuroendocrinology* 32(7): 793-802.

Reese, H. E., McNally, R. J., Najmi, S. and Amir, N. (2010). "Attention training for reducing spider fear in spider-fearful individuals." *J Anxiety Disord*.

Reilly, J. G., McTavish, S. F. and Young, A. H. (1997). "Rapid depletion of plasma tryptophan: a review of studies and experimental methodology." *Journal of Psychopharmacology* 11(4): 381-392.

Rescorla, R. A. and Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. . *Classiacal Conditioning II: Current research and theory*. A. H. Black and W. F. Prokasy. New York, Appleton-Centuary-Crofts**:** 64-99.

Ridout, N., Astell, A. J., Reid, I. C., Glen, T. and O'Carroll, R. E. (2003). "Memory bias for emotional facial expressions in major depression." *Cognition and Emotion* 17(1): 101-122.

Risch, N., Herrell, R., Lehner, T., Liang, K. Y., Eaves, L., Hoh, J., Griem, A., Kovacs, M., Ott, J. and Merikangas, K. R. (2009). "Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: a meta-analysis." *JAMA* 301(23): 2462-71.

Robinson, L. J., Thompson, J. M., Gallagher, P., Goswami, U., Young, A. H., Ferrier, I. N. and Moore, P. B. (2006). "A meta-analysis of cognitive deficits in euthymic patients with bipolar disorder." *Journal of Affective Disorders* 93(1-3): 105-15.

Robinson, O., Cools, R., Crockett, M. and Sahakian, B. (2009). "Mood state moderates the role of serotonin in cognitive biases." *Journal of Psychopharmacology*.

Rolls, E. T. (2004). "The functions of the orbitofrontal cortex." *Brain Cogn.* 55(1): 11-29.

Rudebeck, P. H., Behrens, T. E., Kennerley, S. W., Baxter, M. G., Buckley, M. J., Walton, M. E. and Rushworth, M. F. (2008). "Frontal cortex subregions play distinct roles in choices between actions and stimuli." *J Neurosci* 28(51): 13775-85.

Ruhe, H. G., Mason, N. S. and Schene, A. H. (2007). "Mood is indirectly related to serotonin, norepinephrine and dopamine levels in humans: a meta-analysis of monoamine depletion studies." *Molecular Psychiatry* 12(4): 331-359.

Rushworth, M. F., Mars, R. B. and Summerfield, C. (2009). "General mechanisms for making decisions?" *Curr Opin Neurobiol* 19(1): 75-83.

Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R. and Vuilleumier, P. (2005). "Emotion and attention interactions in social cognition: brain regions involved in processing anger prosody." *Neuroimage* 28(4): 848-58.

Schelling, G., Roozendaal, B. and De Quervain, D. J. (2004). "Can posttraumatic stress disorder be prevented with glucocorticoids?" *Annals of the New York Academy of Science* 1032: 158-166.

Schienle, A., Schafer, A., Hermann, A., Rohrmann, S. and Vaitl, D. (2007). "Symptom provocation and reduction in patients suffering from spider phobia: an fMRI study on exposure therapy." *Eur Arch Psychiatry Clin Neurosci* 257(8): 486-93.

Schmidt, N. B., Richey, J. A., Buckner, J. D. and Timpano, K. R. (2009). "Attention training for generalized social anxiety disorder." *Journal of Abnormal Psychology* 118(1): 5-14.

Schneider, F., Weiss, U., Kessler, C., Müller-Gärtner, H. W., Posse, S., Salloum, J. B., Grodd, W., Himmelmann, F., Gaebel, W. and Birbaumer, N. (1999). "Subcortical correlates of differential classical conditioning of aversive emotional reactions in social phobia." *Biological Psychiatry* 45(7): 863-871.

Schultz, W., Dayan, P. and Montague, P. R. (1997). "A neural substrate of prediction and reward." *Science* 275(5306): 1593-9.

Scott, J., Paykel, E., Morriss, R., Bentall, R., Kinderman, P., Johnson, T., Abbott, R. and Hayhurst, H. (2006). "Cognitive-behavioural therapy for severe and recurrent biploar disorders." *British Journal of Psychiatry* 188(313-320).

See, J., Macleod, C. M. and 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): 65-75.

Seger, C. A. (2006). "The basal ganglia in human learning." *The Neuroscientist* 12(4): 285-290.

Sheline, Y. I., Barch, D. M., Donnelly, J. M., Ollinger, J. M., Snyder, A. Z. and Mintun, M. A. (2001). "Increased amygdala response to masked emotional faces in depressed subjects resolves with antidepressant treatment: An fMRI study." *Biological Psychiatry* 50: 651-658.

Shin, L. M., Rauch, S. L. and Pitman, R. K. (2006). "Amygdala, medial prefrontal cortex, and hippocampal function in PTSD." *Annals of the New York Academy of Science* 1071: 67-79.

Shin, L. M., Whalen, P. J., Pitman, R. K., Bush, G., Macklin, M. L., Lasko, N. B., Orr, S. P., McInerney, S. C. and Rauch, S. L. (2001). "An fMRI study of anterior cingulate function in posttraumatic stress disorder." *Biological Psychiatry* 50(12): 932-942.

Siegle, G. J., Steinhauer, S. R., Thase, M. E., Stenger, V. A. and Carter, C. S. (2002). "Can't shake that feeling: Event-related fMRI assessment of sustained amygdala activity in response to emotional information in depressed individuals." *Biological Psychiatry* 51: 693-707.

Simon, N. M., Otto, M. W., Wisniewski, S. R., Fossey, M., Sagduyu, K., Frank, E., Sachs, G. S., Nierenberg, A. A., Thase, M. E. and Pollack, M. H. (2004). "Anxiety disorder comorbidity in bipolar disorder patients: data from the first 500 participants in the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD)." *American Journal of Psychiatry* 161(12): 2222-2229.

Simon, N. M., Zalta, A. K., Otto, M. W., Ostacher, M. J., Fischmann, D., Chow, C. W., Thompson, E. H., Stevens, J. C., Demopulos, C. M., Nierenberg, A. A. and Pollack, M. H. (2007). "The association of comorbid anxiety disorders with suicide attempts and suicidal ideation in outpatients with bipolar disorder." *Journal of Psychiatric Research* 41(3-4): 255-264.

Smith, K. A., Fairburn, C. G. and Cowen, P. J. (1997). "Relapse of depression after rapid depletion of tryptophan." *The Lancet* 349(9056): 915-919.

Spielberger, C. D., Gorsuch, R. L. and Lushene, R. D. (1983). *Manual for the State-Trait Anxiety Inventory (STAI)*. Palo Alto, CA, Consulting Psychologists Press.

Spitzer, R. L., Williams, J. B. and Gibbon, M. (2002). *Structured Clinical Interview for the DSM-IV*. New York, New York State Psychiatric Institute.

Steel, C., Wykes, T., Ruddle, A., Smith, G., Shah, D. M. and Holmes, E. A. (2010). "Can we harness computerised cognitive bias modification to treat anxiety in schizophrenia? A first step highlighting the role of mental imagery." *Psychiatry Res*.

Storch, E. A., Merlo, L. J., Bengtson, M., Murphy, T. K., Lewis, M. H., Yang, M. C., Jacob, M. L., Larson, M., Hirsh, A., Fernandez, M., Geffken, G. R. and Goodman, W. K. (2007). "D-cycloserine does not enhance exposure-response prevention therapy in obsessive-compulsive disorder." *International Clinical Psychopharmacology* 22(4): 230-7.

Storosum, J. G., Fouwels, A., Gispen-de Wied, C. C., Wohlfarth, T., van Zwieten, B. J. and van den Brink, W. (2004). "How real are patients in placebo-controlled studies of acute manic episode?" *Eur Neuropsychopharmacol* 14(4): 319-23.

Straube, T., Mentzel, H. J. and Miltner, W. H. (2006). "Neural mechanisms of automatic and direct processing of phobogenic stimuli in specific phobia." *Biological Psychiatry* 59(2): 162-170.

Sturmer, B., Leuthold, H., Soetens, E., Schroter, H. and Sommer, W. (2002). "Control over location-based response activation in the Simon task: behavioral and electrophysiological evidence." *J Exp Psychol Hum Percept Perform* 28(6): 1345-63.

Summerfield, C. and Koechlin, E. (2008). "A neural representation of prior information during perceptual inference." *Neuron* 59(2): 336-47.

Sutton, R. and Barto, A. G. (1998). *Reinforcement Learning*. Cambridge, Massachusetts, MIT Press.

Taghavi, M. R., Dalgleish, T., Moradi, A. R., Neshat-Doost, H. T. and Yule, W. (2003). "Selective processing of negative emotional information in children and adolescents with Generalized Anxiety Disorder." *Br J Clin Psychol* 42(Pt 3): 221-30.

Teasdale, J. D. and Dent, J. (1987). "Cognitive vulnerability to depression: an investigation of two hypotheses." *Br J Clin Psychol* 26 ( Pt 2): 113-26.

Tillfors, M., Furmark, T., Marteinsdottir, I. and Fredrikson, M. (2002). "Cerebral blood flow during anticipation of public speaking in social phobia: a PET study." *Biological Psychiatry* 52(11): 1113-1119.

Torres, I. J., Boudreau, V. G. and Yatham, L. N. (2007). "Neuropsychological functioning in euthymic bipolar disorder: a meta-analysis." *Acta Psychiatrica Scandanavia* s434: 17-26.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. and Nelson, C. (2009). "The NimStim set of facial expressions: Judgments from untrained research participants." *Psychiatry Res*.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J. and Nelson, C. (2009). "The NimStim set of facial expressions: judgments from untrained research participants." *Psychiatry Res* 168(3): 242-9.

Tranter, R., Bell, D., Gutting, P., Harmer, C. J., Healy, D. and Anderson, I. M. (2009). "The effect of serotonergic and noradrenergic antidepressants on face emotion processing in depressed patients." *J Affect Disord* 118(1-3): 87-93.

Van Damme, S., Legrain, V., Vogt, J. and Crombez, G. (2010). "Keeping pain in mind: a motivational account of attention to pain." *Neurosci Biobehav Rev* 34(2): 204-13.

van der Veen, F. M., Evers, E. A., Deutz, N. E. and Schmitt, J. A. (2007). "Effects of acute tryptophan depletion on mood and facial emotion perception related brain activation and performance in healthy women with and without a family history of depression." *Neuropsychopharmacology* 32(1): 216-224.

van Honk, J., Peper, J. S. and Schutter, D. J. (2005). "Testosterone reduces unconscious fear but not consciously experienced anxiety: implications for the disorders of fear and anxiety." *Biological Psychiatry* 58: 218-225.

Velten, E., Jr. (1968). "A laboratory task for induction of mood states." *Behav Res Ther* 6(4): 473-82.

Vuilleumier, P. (2005). "How brains beware: neural mechanisms of emotional attention." *Trends in Cognitive Science* 9(12): 585-594.

Vuilleumier, P. and Driver, J. (2007). "Modulation of visual processing by attention and emotion: windows on causal interactions between human brain regions." *Philos Trans R Soc Lond B Biol Sci* 362(1481): 837-55.

Wadlinger, H. A. and Isaacowitz, D. M. (2008). "Looking happy: the experimental manipulation of a positive visual attention bias." *Emotion* 8(1): 121-126.

Walton, M. E., Behrens, T. E., Buckley, M. J., Rudebeck, P. H. and Rushworth, M. F. (2010). "Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning." *Neuron* 65(6): 927-39.

Watkins, E., Moberly, N. J. and Moulds, M. L. (2008). "Processing mode causally influences emotional reactivity: distinct effects of abstract versus concrete construal on emotional response." *Emotion* 8(3): 364-78.

Watson, D., Clark, L. A. and Tellegen, A. (1988). "Development and validation of brief measures of positive and negative affect: the PANAS scales." *J Pers Soc Psychol.* 54(6): 1063-1070.

Wells, T. T. and Beevers, C. G. (2009) "Biased attention and dysphoria: Manipulating selective attention reduces subsequent depressive symptoms." *Cognition and Emotion*, 10.1080/02699930802652388: 10.1080/02699930802652388

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B. and Jenike, M. A. (1998). "Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge." *Journal of Neuroscience* 18(1): 411-418.

Wilhelm, S., Buhlmann, U., Tolin, D. F., Meunier, S. A., Pearlson, G. D., Reese, H. E., Cannistraro, P., Jenike, M. A. and Rauch, S. L. (2008). "Augmentation of behavior therapy with D-cycloserine for obsessive-compulsive disorder." *American Journal of Psychiatry* 165(3): 335-41; quiz 409.

Williams, J. M. G., Mathews, A. and MacLeod, C. (1996). "The emotional Stroop task and psychopathology." *Psychological Bulletin* 120(1): 3-24.

Williams, J. M. G., Watts, F. N., MacLeod, C. M. and Mathews, A. (1997). *Cognitive Psychology and Emotional Disorders*. New York, John Wiley & Sons.

Wilson, E. and MacLeod, C. M. (2003). "Contrasting two accounts of anxiety-linked attentional bias: Selective attention to varying levels of stimulus threat intensity." *Journal of Abnormal Psychology* 112(2): 212-218.

Worsley, K. J., Evans, A. C., Marrett, S. and Neelin, P. (1992). "A three-dimensional statistical analysis for CBF activation studies in human brain." *Journal of Cerebral Blood Flow and Metabolism* 12(6): 900-918.

Yiend, J. and Mathews, A. (2001). "Anxiety and attention to threatening pictures." *Quarterly Journal of Experimental Psychology A* 54(3): 665-681.

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A. and Perrett, D. I. (1997). "Facial expression megamix: tests of dimensional and category accounts of emotion recognition." *Cognition* 63: 271–313.

Young, R. C., Biggs, J. T., Ziegler, V. E. and Meyer, D. A. (1978). "A rating scale for mania: reliability, validity and sensitivity." *British Journal of Psychiatry* 133: 429-435.