Measuring and Modifying Information Bias in Depression



Sungwon Han

Saint Hilda's College, University of Oxford

Thesis submitted for the degree of Master in Science (by Research) in Psychiatry

Trinity Term 2017

ACKNOWLEDGEMENTS

I would first and foremost like to thank my supervisor Dr. Michael Browning for his warm support throughout the completion of my project. I also wish to express my gratitude to Dr. Erdem Pulcu for teaching me coding. With these academic advisors, I was able to complete this thesis and develop my academic ability.

I especially thank to members of St Hilda's College. I am grateful to Dr. Stephen McHugh, who is my college advisor, for advising me in various academic discussions, which helped me have a broad view as a potential scientist, and to Middle Common Room of the College for their unceasing support for postgraduate students.

I appreciate the government of the Republic of Korea for their financial support for my study in Oxford. My work in current Master's degree programme is funded by the Korean Government Scholarship Program for Study Overseas (*gukbiyuhak hapgyeok* No. 2015-74 – *gwalli* No. 2015-79; 대한민국 정부 국비유학 합격 2015-74호 – 관리 2015-79호). Director Tae-il Kim, Korean Education Centre in the UK, and Education Researcher Eunsuk Kim, National Institute for International Education of Ministry of Education in Korea, kindly supported my study in administrative ways. Moreover, I also appreciate for Mr. Edward Kwanhun Rim for his financial support for my academic expenditure.

Most of all, I sincerely appreciate all participants for their interest and cooperation for the study. Their precious time and understanding on the study are always the biggest supports for our researchers.

Neurosciences Building at Warneford Hospital, August 2017

ABSTRACT

Measuring and Modifying Information Bias in Depression

Sungwon Han

St. Hilda's College, University of Oxford

Thesis submitted for the degree of Master of Science (by Research) in Psychiatry

Trinity Term 2017

Background: Depression is a common illness which is believed to be caused, at least in part, by a tendency to preferentially process negative, at the expense of positive, events. Recent work in computational neuroscience indicates that individuals estimate how informative an event is and are more influenced by those events they judge to be more informative. This suggests a potential mechanism which may lead an individual to develop a negative bias—they may estimate that negative events are more informative than positive events. It also suggests a novel approach to the treatment of depression—interventions which reduce the estimated information content of negative relative to positive events should act to improve mood. A number of factors can influence how informative an event is, one of which is the changeability, or volatility, with which it is associated with a given outcome. The main aim of this study was to examine the role of the estimated information content of positive and negative outcomes, in depression. The study addresses two specific questions: First, do previously depressed patients, relative to never depressed controls, demonstrate a tendency to over-estimate the information content of negative outcomes? Second, can a simple training intervention alter an individual's estimate of the information content of positive relative to negative outcomes, and does this influence other cognitive measures and symptoms of depression? The study described in this thesis is a large ongoing clinical study. The thesis describes an analysis of data from this study collected by myself during the first year of its recruitment, data collection for the larger study is ongoing. *Methods*: The study assessed the baseline cognitive biases using a volatility learning task and standard face based task, as well as depressive symptoms in two groups of participants: Previously depressed patients (N=16) and never depressed controls (N=18). After this, the participants were randomly assigned into either positive or neutral learning training for two weeks, and were tested using the same outcome measures immediately after training and then again four weeks later. *Results*: Across groups participants used a higher learning rate for informative relative to non-informative outcomes indicating that participants' behaviour was influenced by the information content of the outcomes. No baseline difference in learning rate and inverse temperature between groups was found, although previously depressed patients did show a negative bias on the faces task. The study found that the learning training intervention led to the expected alteration in learning rates as well as having some impact on measures of anhedonia. *Conclusion*: Modification of learning rate for positive relative to negative outcomes with training was achieved although it did not lead to a convincing antidepressant effect. The facial based task is an efficient tool to measure previously depressed individuals' negative cognitive bias. It will be important to assess the impact of learning training in a larger sample of participants.

CONTENTS

Ⅰ. Introduction 10

1. The Cost and Impact of Depression 10

2. Cognitive Approach for Understanding Depression 13

3. Altering Bias, as a Potential Approach to the Treatment of Depression 17

4. Computational Accounts of Why People Might Show These Biases 19

5. The Current Study 25

Ⅱ. Materials and Methods 26

1. Study Design 26

2. My Role in the Study 26

3. Participants 27

4. Procedure 27

5. Questionnaire Measures 30

6. Facial Expression Recognition Task (FERT) 33

7. Information Bias Learning Task (IBLT) & Information Bias Assessment Task (IBAT) 35

8. Computational Model 38

9. ANOVA & T-tests Analysis 40

Ⅲ. Result 42

1. Sample Demographics 42

2. IBAT Task: Learning Rate 44

3. IBAT Task: Inverse Temperature 49

4. FERT 55

4.1. FERT Task: Baseline Comparison 58

4.2. The Effect of IBLT on the FERT Task 64

5. Questionnaire Measures 71

6. QIDS 71

7. SHAPS 72

8. Trait-STAI 73

9. DAQ 74

10. RSS 75

Ⅳ. Discussion 76

REFERENCE 85

APPENDICES 99

LIST OF TABLES

Table 1. Inclusion and Exclusion Criteria of the Study 27

Table 2. Baseline Characteristics 43

Table 3. FERT Accuracy Changes 55

Table 4. FERT Misclassification Changes 56

Table 5. FERT Reaction Time Changes 57

LIST OF FIGURES

Figure 1. Global Mean Health by Disease Status in 2003 (Moussavi et al., 2007, p. 854) 12

Figure 2. Summary of Learning Task with Volatile & Stable Outcome Probabilities (Pulcu & Browning, 2017, DOI: 10.7554/eLife.27879.002) 23

Figure 3. Flow Chart of the Study 28

Figure 4. One Trial and Possible Outcomes of the Information Bias Tasks (Both IBAT & IBLT) 36

Figure 5. Structure of the Information Bias Assessment Task (IBAT) 36

Figure 6. Average Points of Daily Won During IBLT by 4 Groups 44

Figure 7. Learning Rates (LR) of the Two Participant Groups at Baseline 45

Figure 8. The Effect of IBLT upon Learning Rate over Time 48

Figure 9. Inverse Temperatures (IT) of the Two Participant Groups at Baseline 49

Figure 10. Effect of Neutral IBLT upon Inverse Temperature in the Previously Depressed Group for “Win” & “Loss” Outcomes over Time 52

Figure 11. Effect of Positive IBLT upon Inverse Temperature in the Previously Depressed Group for “Win” & “Loss” Outcomes over Time 52

Figure 12. Effect of Neutral IBLT upon Inverse Temperature in the Control Group for “Win”

& “Loss” Outcomes over Time 54

Figure 13. Effect of Positive IBLT upon Inverse Temperature in the Control Group for “Win” & “Loss” Outcomes over Time 54

Figure 14. Baseline Accuracy in Negative & Positive Emotions 59

Figure 15. Baseline Accuracy in 6 Individual Emotions 59

Figure 16. Baseline Misclassification in Negative & Positive Emotions 61

Figure 17. Baseline Misclassification in 6 Individual Emotions 61

Figure 18. Baseline Reaction Time in Negative & Positive Emotions 63

Figure 19. Baseline Reaction Time in 6 Individual Emotions 63

Figure 20. Effect of IBLT upon Accuracy over Time in Negative Emotions 65

Figure 21. Effect of IBLT upon Accuracy over Time in Positive Emotions 65

Figure 22. Effect of IBLT upon Accuracy over Time in Individual Emotions 65

Figure 23. Relative Accuracy of the Neutral & Positive IBLT Groups for the 6 Emotions of the FERT Task 66

Figure 24. Effect of IBLT upon Misclassification over Time in 6 Individual Emotions 67

Figure 25. Effect of IBLT upon Misclassification over Time in Negative Emotions 67

Figure 26. Effect of IBLT upon Misclassification over Time in Positive Emotions 67

Figure 27. Interaction between Emotional Valence and Subject Group (Reaction Time) 68

Figure 28. Effect of IBLT upon Reaction Time over Time in Negative Emotions 69

Figure 29. Effect of IBLT upon Reaction Time over Time in Positive Emotions 69

Figure 30. Interaction between 6 Individual Emotions & Subject Group (Reaction Time) 70

Figure 31. Effect of IBLT upon Reaction Time over Time in Negative & Positive Emotions 70

Figure 32. Effect of IBLT upon QIDS over Time 71

Figure 33. Effect of IBLT upon SHAPS over Time 72

Figure 34. Effect of IBLT upon Trait-STAI over Time 73

Figure 35. Effect of IBLT upon DAQ over Time 74

Figure 36. Effect of IBLT upon RSS over Time 75

LIST OF ABBREVIATIONS

ABM	Attentional Bias Modification
ANOVA	Analysis of Variance
CAR	Cortisol Awaking Response
CBM	Cognitive Bias Modification
CBT	Cognitive Behavioural Therapy
COPD	Chronic Obstructive Pulmonary Disease
DAQ	Depressive Attributions Questionnaire
DSM-IV (DSM-4)	Diagnostic and Statistical Manual of Mental Disorders - Fourth Edition
FERT	Facial Expression Recognition Task
IBAT	Information Bias Assessment Task
IBLT	Information Bias Learning Task
IDS- C_{16}	Inventory of Depressive Symptomatology – Clinician Scale
IDS- SR_{16}	Inventory of Depressive Symptomatology – Self-Rating Scale
IT	Inverse Temperature
LR	Learning Rate
RRS	Ruminative Response Scale
SCID-V (SCID-5)	Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders
SD	Standard Deviation
SHAPS	Snaith–Hamilton Pleasure Scale
STAI	State-Trait Anxiety Inventory
M	Mean
MDD	Major Depressive Disorder
MDE	Major Depressive Episodes
MADRS	Montgomery–Åsberg Depression Rating Scale
PIS	Participant Information Sheet
YLDs	Years Lived with Disability
QIDS- C_{16}	Quick Inventory of Depressive Symptomatology – Clinician Scale
QIDS- SR_{16}	Quick Inventory of Depressive Symptomatology – Self-Rating Scale

Ⅰ. Introduction

1. The Cost and Impact of Depression

Depression is a serious and common psychiatric illness with a high prevalence and large impact in terms of life quality and economic cost. *Prevalence:* The international epidemiological study of major depressive episodes (MDE) in 2003 investigated the prevalence of depression in ten nations in different time settings (Andrade et al., 2003). According to the study, the lifetime prevalence of MDE ranged between 3% (Japan) and 16.9% (US), where all national surveys except for the Japanese one were over 5% (Andrade et al., 2003). Twelve-month prevalence was in the range between 1.2% (Japan) to 10% (US) (Andrade et al., 2003, p. 9). Thirty-day prevalence resulted in between 0.9% (Japan) and 4.6% (US) (Andrade et al., 2003, p. 9). In another study in England in 2000, approximately 2.6 million cases of depression were reported, many of whom were of working age (70.60% of cases were between the age of 25 and 64) (Thomas & Morris, 2003, p.515). *Life Quality:* In a study of primary care patients, both anxiety and depression were reported to be significantly associated with quality of life (Brenes, 2007, p. 438-440). A six-month follow-up study reported that remitted depressed patients who left hospital showed better quality of life than currently depressed patients did, but worse quality of life than non-depressed people did. (Angermeyer, Holzinger, Matschinger & Stengler-Wenzke, 2002, p. 192-196). Globally, major depressive disorder (MDD) is ranked as the second leading cause of years lived with disability (YLDs) (Vos et al., 2015, p.765). Unsurprisingly the functional impact of depression (with or without comorbidities) is substantial with a marked negative

impact on social functioning (Hecht, von Zerssen & Wittchen, 1990, p. 140-142). In terms of mental and physical health-related quality of life, the treatment of depression has been studied to see whether it is associated with or contributes to the life quality, meaning that as depression is treated, the patient gets better health-related quality of life. Although limited to patients with primary cardiac illness, a meta-analysis found that the treatment of depression is effective at improving mental health related life quality (O'neil, Sanderson, Oldenburg & Taylor, 2011, p. 153-155). *Financial Burden:* The total cost of depression in England in 2000, summing up primary, secondary and societal care costs, was evaluated as approximately 9.06 billion pounds (Thomas & Morris, 2003, p.516).

Comorbid Depression & Physical Illnesses: Depression is commonly comorbid with other psychiatric and medical illness, and is often associated with poorer outcomes when it is found in such cases. For example, depression is associated with a significant decrement in the health of people with chronic diseases such as angina, arthritis, asthma, diabetes and chronic obstructive pulmonary disease (COPD) (Moussavi, Chatterji, Verdes, Tandon, Patel & Ustun, 2007). The researchers analysed the health of people in 60 countries in order to compare the impact of depression on patients with chronic diseases. Significant numbers of respondents with chronic diseases answered that they also have depression (Moussavi et al., 2007, p. 854). In addition, mean health score analysis reported that those who had “depression and two/more chronic conditions” showed the lowest mean health score (56.1), as well as that all scores of those who had depression and each of the chronic diseases (58.5 to 67.1) were lower than scores of those who had "no chronic condition" (90.6), each chronic disease

(78.9 to 80.3), "depression only" (72.9), and "two/more chronic conditions" (71.8) (Moussavi et al., 2007, p. 854, *Figure 1* below).

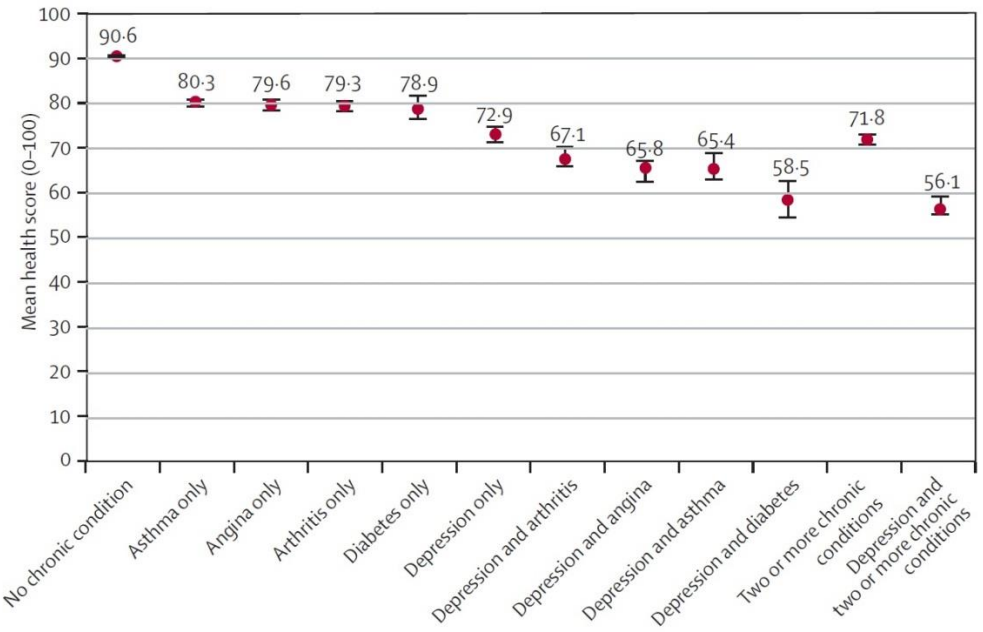


***Figure 1*. Global Mean Health by Disease Status in 2003 (Moussavi et al., 2007, p. 854)**

Treatment Efficacy in Depression: Another problem faced by depressed patients is that the efficacy of available treatments is relatively modest. Casacalenda, Perry & Looper (2002) assessed the remission of depression in patients treated with pharmacological or psychological interventions as well as in an untreated control condition from six previous studies, whose treatment periods were between 10 weeks and 8 months. According to the analysis, remission rates of active treatments (pharmacology 46.4%, psychotherapy 46.3%) were superior to that of controls (24.4%) (Casacalenda et al., 2002, p. 1357), but as can be seen, the absolute rate of remission in both classes of active treatment was less than half of those treated. Many previous studies also have investigated recurrence in

depression. Keller, Shapiro, Lavori & Wolfe (1982) followed recovery of 75 depressed patients for 12 months. While the probability of recurrence reduced over time, 25% had a recurrence within 12 weeks and 12% had one within just 4 weeks at risk (Keller et al., 1982, p. 912-914). The frequent recurrence of depression after recovery has been reported in previous findings (Keller, Lavori, Lewis & Klerman, 1983; Gonzales, Lewinsohn & Clarke, 1985; Roy-Byrne, Post, Uhde, Porcu & Davis, 1985; Frank et al., 1990; Maj, Veltro, Pirozzi, Lobrace & Magliano, 1992).

In summary, depression is a common condition affecting many people worldwide, and is associated with significant personal, medical and financial costs. In addition, depression is very commonly recurrent and has treatments of only modest efficacy. There is clearly a strong need for more effective therapies for this illness: an essential prerequisite for this is a better understanding of the underlying processes which lead certain individuals to develop depression.

2. Cognitive Approach for Understanding Depression

Cognitive models suggest that difficulties or habits in the way that individuals process information may contribute to the onset or maintenance of the mental disorders, such as panic disorder, social phobia, post-traumatic stress disorder, generalised anxiety disorder and depression (Mathews & Macleod, 2005, p. 168, 179-188).

The cognitive literature in depression can be roughly split into studies which

report some form of underperformance on measures of cognition (called here “a deficit”) and those which report some differential preference for one from of information over another (called here “a bias”).

Deficit Studies: Previous investigators have reported that depression is associated with cognitive deficits, with the severity of the deficits being linked to the severity and chronicity of the illness. First, meta-analyses indicate that depressed patients show, relative to control subjects, poorer task performance on measures of episodic memory, executive function and processing speed: with cognitive performance being more impaired in patients with more severe depression (McDermott & Ebmeier, 2009, p. 4). Increasing cognitive deficit has also been associated with increasing chronicity of depression (Dotson, Resnick & Zonderman, 2008, p. 322-329). As a specific example of the primary data which contributes to the meta-analysis described above, depressed people, compared to controls, show lower accuracy, longer movement time and longer thinking time for a spatial planning task (the Tower of London) (Beats, Sahakian & Levy, 1996, p. 597). A recent review has also suggested that remitted depressed patients show significantly poorer cognitive function than people who have never been depressed (Rock, Roiser, Riedel & Blackwell, 2014, p. 2034). Other researchers also have shown cognitive deficits of depression in particular measures of memory recall (Gorwood, Corruble, Falissard & Goodwin, 2008, p. 733-738) and prefrontal function (Mayberg, 2002, p. 376-391).

Bias Studies: Bias studies have investigated the tendency for patients with depression (or at risk of depression) to preferentially process certain forms of information at the expense of others. More specifically, depressed patients will

tend to prioritise negative at the expense of positive information during processing, a phenomenon termed “negative affective bias”. Negative cognitive bias has been reported in studies of memory (Matt, Vázquez & Campbell, 1992), attention (Gotlib, Krasnoperova, Yue & Joormann, 2004, p. 131-132) and interpretation (Rude, Wenzlaff, Gibbs, Vane & Whitney, 2002, p. 429-434). For example, MDD individuals showed 10% better memory for negative words than positive ones (Matt et al., 1992, p. 238, 242). Moreover, comparing between depressed and healthy individuals, depressed patients had better memory (both explicit & implicit) for negative material, whereas non-depressed individuals tended to show better memory performance for positive material (Rinck & Becker, 2005, p. 69-70). However, the literature on bias measures does show some variability with previous studies reporting no deficits of depressive individuals in implicit memory using repetition primary and word-fragment completion paradigms (Danion, Willard-Schroeder, Zimmermann, Grangé, Schlienger, & Singer, 1991, p. 709; Jenkins & McDowall, 2001, p. 809), with other studies reporting a worse performance by depressed people than control participants in a neutral implicit memory task using word stem completion, homophone spelling and category association paradigms (Elliott & Greene, 1992, p. 573-574; Jenkins & McDowall, 2001, p. 809-810).

A second approach to the question of bias is to study how patients learn from positive versus negative outcomes. Previous studies have reported that depressed patients show deficits in the processing of rewards, including during reinforcement learning and decision making. For example, depressed patients rated positive stimuli as less pleasant and less arousing (Sloan, Strauss, Quirk & Sajatovic, 1997, p. 138-139), and showed reduced pleasurable response to

positive stimuli (Sloan, Strauss & Wisner, 2001, p. 490-492), although a counter example indicates that depressed and control subjects do not differ in the sweet taste test (Dichter, Smoski, Kampov-Polevoy, Gallop & Garbutt, 2010, p. 861-862). McFarland and Klein (2009, p. 119-120) also found out that depressed individuals were less sensitive than control participants when anticipating reward, with no difference for anticipating punishment. Individuals with MDD have also been reported to show less response to rewarded stimulus (Pizzagalli, Iosifescu, Hallett, Ratner & Fava, 2008, p. 80-83). This feature, depressed people's hyposensitivity to reward, was also seen in a study of working memory (Christopher & MacDonald, 2005, p. 388-392) and using a signal detection task (Pizzagalli, Jahn & O'Shea, 2005, p. 322-324). A meta-analysis of studies employing the same signal detection task also suggested that depressed patients and those with anhedonia displayed hyposensitivity to reward (Huys, Pizzagalli, Bogdan & Dayan, 2013, p. 8-9). Altered response to rewards and punishments can also be measured in learning tasks. For example, Must, Szabó, Bódi, Szász, Janka and Kéri (2006, p. 212-213) reported that people with MDD are less able to learn the best choice to make in a reward task but have preserved function in a punishment task relative to a non-depressed group. Moreover, reduced depressive symptoms were associated with patients' being relatively less influenced by punishment (Must et al., 2006, p. 213).

Lastly, the bias for negative information in depressed patients has been described in terms of patients' response to task feedback, with depressed patients being described as experiencing an extreme response (Elliott, Sahakian, Herrod, Robbins & Paykel, 1997, p. 76-79). Specifically, it is argued that negative feedback has a disproportionate effect on the decisions of depressed patients, such that

they perform tasks sub-optimally.

In summary, despite a degree of variability in the published literature, there is evidence for both cognitive deficits and negative affective bias in patients with depression and those who are at risk of developing depression. In this thesis I will focus on biased processing, why it develops and how it may best be altered by treatments.

3. Altering Bias, as a Potential Approach to the Treatment of Depression

The research reviewed above describes cognitive phenomena, such as deficits or biases, which are associated with depression. However, processes which are causally related to depression, rather than simply being associated with it, are of particular interest as they are then potential candidates for treatment targets (*N.B.* a treatment target is a process which is altered by treatments, leading to an improvement in symptoms). A number of lines of evidence support the role of cognitive bias as a causal risk factor, and therefore treatment target, in depression. Firstly, psychological interventions such as Cognitive Behavioural Therapy (CBT) were explicitly designed to, among other things, alter automatic negative thoughts, cognitive distortions, and ultimately negative schemata processing bias (Beck, 1967, p. 318-319; Beck, 1976, p. 264-267, 301; Beck, Rush, Shaw & Emery, 1979, p. 35). This theory postulates that the negative cognitive set can be modified by positive emotional images, and this will reduce negative thinking and emotion (“cognitive modification”) (Beck, 1976, p. 267-269). The efficacy of this form of therapy provides some support for the proposal that negative biases are causally

related to depression (Driessen & Hollon, 2010, p. 547-550). Secondly, researchers have postulated that antidepressants work, at least in part, by altering emotional bias (Pringle & Harmer, 2015, p. 477-478, 484). Specifically, many pharmacological treatments for depression have been found to reduce negative, or increase positive, bias (Harmer et al., 2009, p. 1178-1179). This work has shown that antidepressants alter processing bias in depressed patients before clinical symptoms of depression improve, and that the early effect of the treatment on affective bias predicts later clinical response. In Tranter and colleagues' investigation of how antidepressants effect face emotion processing, antidepressant treatment altered the recognition of happy expressions (Tranter et al., 2009) after 2 weeks of treatment. Crucially, the degree to which this positive bias occurred at 2 weeks predicted the extent to which patients' symptoms of depression responded after a full 6 weeks of treatment. Both of these observations support a causal role for affective bias in the symptoms of depression. Lastly there is a body of work which examines techniques specifically designed to alter the affective biases using simple training tasks. These so called "cognitive bias modification (CBM)" tasks have been shown to have some effect on symptoms of depression (Browning, Holmes & Harmer, 2010, p. 9, 14-17; Browning, Holmes, Charles, Cowen & Harmer, 2012, p. 575-577), although questions have been raised about the size and reliability of this effect (Hallion & Ruscio, 2011, 949-954; Cristea, Kok & Cuijpers, 2015, p. 9-15). As an example of the CBM literature, Browning et al. (2012, p. 573-577) tested whether using attentional bias modification (ABM) to alter negative attentional bias in previously depressed people affected markers of the probability of recurrence of depression, and

showed that positive ABM (which induces a relatively positive bias in attentional function) reduced both residual symptoms and cortisol waking response, both of which are associated with risk of relapse. In other words, altering attentional bias in an experimental context influenced markers of risk of depression, providing relatively strong empirical evidence for the causal role of attentional bias in the illness. Other studies have also investigated the impact of CBM in social anxiety (Liu, Li, Han & Liu, 2017, p. 9-14) and anhedonia (Blackwell et al. 2015, p. 96-103).

4. Computational Accounts of Why People Might Show These Biases

Computational approaches are increasingly being used in the analysis of behaviour and the identification of the cognitive processes which underlie behaviour. Broadly, these approaches seek to emulate a cognitive process, or a set of processes, mathematically. For example, computational analyses of learning tasks in which participants have to learn the values of various options and then select an option based on this learning often include a learning component, in which a simple mathematical learning rule is used with the output of this learning rule then being fed into a decision process, such as a softmax action selector (Reverdy & Leonard, 2016, p. 54). This approach has a number of advantages: Firstly, it exactly specifies the hypothesised form of the cognitive process in mathematical terms, which allows precise testing of the degree to which a model can, or cannot, account for behaviour. Secondly, it is particularly useful at describing dynamic aspects of cognition which change over time. For example, during learning a person's belief about the topic being learned will change every

time new information is presented—this process is more easily captured with a trial based model than with the simple summary metrics (e.g. average accuracy) often used in traditional analyses. Thirdly, computational modelling approaches have been successful at linking together different levels of analysis of a process such as behaviour and cognition to underlying neural systems which produce this behaviour. The most prominent example of this being the finding that the activity of mid brain dopaminergic neurons are well predicted by the reward prediction error term of the temporal difference learning algorithm (Schultz, Dayan & Montague, 1997, p. 1593-1957). Finally, computational approaches can be useful as they may suggest novel ways to address existing questions. For example, the question as to why people may show negative biases, described above, can be addressed from a computational learning perspective. I set out the background to this below.

When learning about the best actions to take or to avoid, people need to deal with a number of sources of uncertainty. Two forms of uncertainty are commonly described as “expected uncertainty” and “unexpected uncertainty” (Yu & Dayan, 2005). The way to distinguish these two kinds of uncertainty is by questioning whether the uncertainty of the event arises because of the basic unpredictability of a familiar environment (expected uncertainty) or is caused by a fundamental shift in the association which is being learned (unexpected uncertainty) (Yu & Dayan, 2005, p. 681). For instance, when a person rolls a dice, he/she knows the probability of each number (1/6) but does not know exactly what the result of each roll will be. This fundamental uncertainty found in all probabilistic situations is called expected uncertainty. Unexpected uncertainty occurs if the dice were swapped to a new dice which was not fair (i.e. the probabilities of the

outcomes were not all 1/6) so he/she had to learn the new outcome probabilities. These forms of uncertainty influence how useful, or informative, an event is during learning. For example, learning about a probabilistic association between an action and an outcome requires that what one already knows about the association is combined with evidence from new experience on an ongoing basis. If the unexpected uncertainty is high, that is if the association itself is very volatile, then previous estimates become less useful (because the association is more likely to have changed since these estimates were formed) and the evidence from new events becomes relatively more useful. In other words, when volatility is high, new events are more informative and thus learners should be more influenced by them than when the volatility is low. This balance between the influence of previous estimates and new events during learning is captured in simple learning algorithms as the learning rate. For instance, the Rescorla-Wagner Model (Rescorla & Wagner, 1972) can describe mathematically how the belief about the event should be updated:

$$r_{(t+1)} = r_{(t)} + \alpha(s_{(t)} - r_{(t)})$$

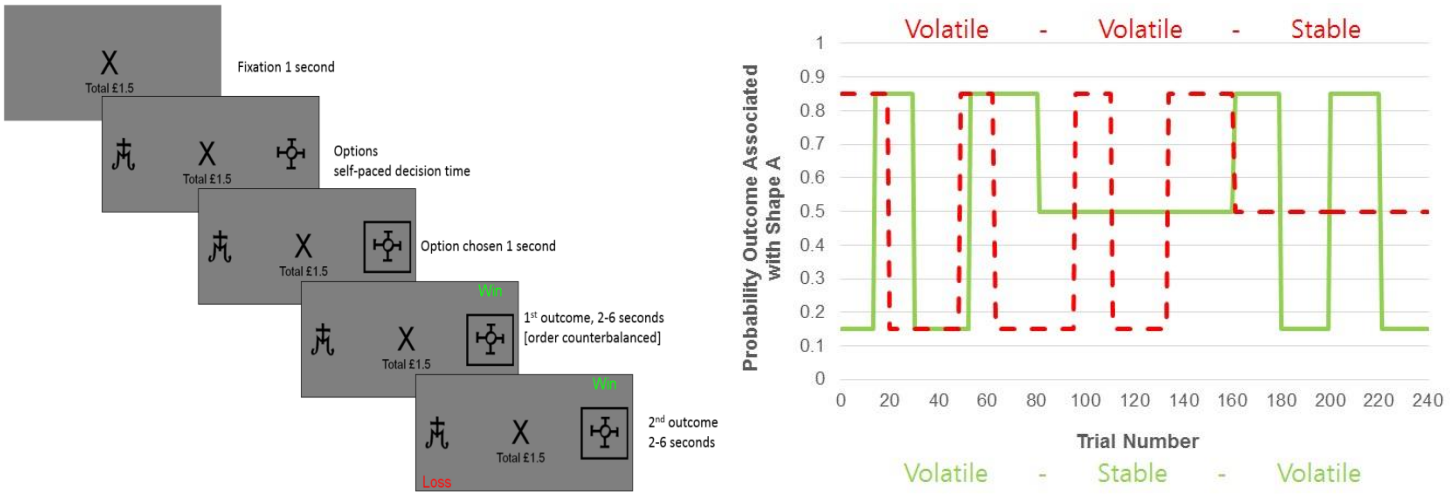
In this equation, $r_{(t)}$ indicates the estimate of the probability of the event at time *t*, and $r_{(t+1)}$ is the updated estimation after the new event, with the update being the learning rate α, multiplied by the prediction error, the difference between the outcome $s_{(t)}$ and the estimate $r_{(t)}$.

A low learning rate means that a learner is putting more weight on previous estimates of an association and is less influenced by the new event, whereas a high learning rate means the learner is putting more weight on new events. As

explained above, when unexpected uncertainty, or volatility, is high learning should be more influenced by recent events, which equates to using a high learning rate. Previous studies which examine learning in humans have consistently found that a higher learning rate is indeed used when volatility is high (Behrens, Woolrich, Walton & Rushworth, 2007, p. 1215-1220; Behrens, Hunt, Woolrich & Rushworth, 2008, p. 245-246, 248; Nassar, Rumsey, Wilson, Parikh, Heasly & Gold, 2012, p. 1040-1045; Browning, Behrens, Jocham, O'Reilly & Bishop, 2015, p. 2-6). In other words, humans tune their learning rates so that they are more influenced by events which are more informative, such as those which occur in a volatile context.

This computational work provides a background for understanding why people should (and when they should) display negative cognitive biases: negative biases are an appropriate response if one estimates the information content of negative events to be higher than for positive events. Equivalently, this suggests that depressed patients are interacting with their world as if the negative events they encounter are more informative than the positive events. In essence the above argument provides a reformulation of the common observation that depressed patients focus more on negative events. This reformulation may have utility if it suggests a novel method of engaging with and altering the biases which are causally related to depression. Again, based on the argument advanced above, it is predicted that it may be possible to alter affective biases by altering an individual’s estimated information content for positive and negative events.

(a) (b)



(c)

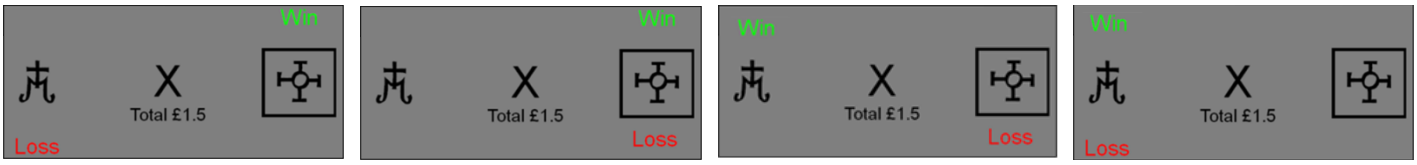


***Figure 2*. Summary of Learning Task with Volatile & Stable Outcome Probabilities (Pulcu & Browning, 2017, DOI: 10.7554/eLife.27879.002)**

The sub-figure (a) is one trial of the learning task that the authors used. Participants were presented with two different shapes on the left and right side and asked choose one of them. On every trial, the two outcomes "win (giving 15 pence as reward)" and "loss (taking 15 pence out as punishment)" were associated with one of the two shapes, and these outcomes were presented after the participants choose a shape at each trial. Over the course of a number of trials, participants learned which shape is more likely to be associated with "win" and "loss" respectively. The sub-figure (b) is the structure of probabilities of "win" and "loss". The task consists of total 240 trials, which can be divided into three blocks (80 trials at each block). The x-axis is the number of trials, and the y-axis is the probability (*p*) of "win" (represented by green line) and "loss" (represented by red line) that are associated with a certain shape. The probability of "win" and "loss" that are associated with the other shape is 1-*p*. At the 1st block, both "win" and "loss" outcomes were volatile (changeable). At the 2nd block, the "win" outcome was stable, but the "loss" one was volatile, and these two are opposite in the 3rd block. This different volatility of two outcomes affect how informative that each outcome is: In the 2nd block, the probability of the "win" outcome is stable which means that this outcome gives less information, but the probability of the "loss" outcome is volatile which means that this outcome provides more information. Therefore, people in this block are expected to learn more about the "loss" outcome than the "win" outcome. The sub-figure (c) shows the four possible outcomes in a trial. The outcomes of "win" and "loss" were independent from each other, therefore, in order to complete the task participants had to separately learn the position of these outcomes separately.

Recent work (Pulcu & Browning, 2017) has investigated whether it is possible to induce this effect in non-clinical volunteers. In this work, participants completed a learning task in which their choices led to both positive (winning money) and negative (losing money) outcomes (*Figure 2*). These outcomes were independent, meaning that a given choice could result in a win, a loss, both a win and a loss or neither a win nor a loss (*Figure 2c*). From the participants' point of view this means that, in order to do the task participants had to learn, separately, where they thought the win would occur and where they thought the loss would occur. The task was divided into three blocks with the volatility of the outcomes manipulated independently (*Figure 2b*). This allows evaluation of the degree to which participants' tuned their learning rates to the information content of the positive ("win") and negative ("loss") outcomes (*NB*: As explained above, a volatile outcome is more informative than a stable outcome.). The results from this study indicated that participants adjusted their learning rate to the information of the two outcomes and that the learning rates were affected by the manipulated volatility. In other words, participants used a higher learning rate for either the "win" or "loss" outcome if that outcome carried more information.

The current study extends this work by assessing whether previously depressed participants differ from never depressed controls using a similar measure of learning and by providing an initial assessment of whether using this volatility based manipulation to alter bias in currently and previously depressed patients lead to changes in mood as well as in other cognitive and biochemical processes.

5. The Current Study

Objectives: Based upon the previous research discussed above, the aim of this study is to use a novel, computationally informed learning task to measure information bias in depressed and control participants (*[1]) and to examine the impact of modifying this bias using a simple training task. Specifically, the current study assessed whether a task which increases learning rate for positive relative to negative outcomes reduces symptoms of depression, and induces a more positive bias in cognitive measures. In other words, the study assessed group differences in learning from positive and negative events, and the effect of a novel learning training intervention on cognitive and mood measures. The study has three hypotheses, described below.

Hypotheses

(1) Depressed people will show a higher learning rate for negative and a lower learning rate for positive events than controls do.

(2) The learning rate of those who took positive learning training (positive IBLT) will be modified to be more positive than the learning rates of those who took neutral learning training (neutral IBLT).

(3) The positive learning training will, relative to the control training, lead to positive changes in mood questionnaires and the Facial Expression Recognition Task (FERT).

[1] Note the larger study is planned to recruit over approximately two years with three groups: Currently depressed, previously depressed and control participants. This thesis analysed interim data, collected by me between November 2016 and May 2017, of only previously depressed and control groups.

Ⅱ. Materials and Methods

1. Study Design

The study used a double-blind randomised design with three groups: (1) “currently depressed” group, (2) “previously depressed but remitted at the moment of participation” group, and (3) “not depressed” group (control group). Differences among groups were assessed in a baseline session, with the effects of a novel intervention (IBLT) assessed by randomizing all participants to receive a positive or neutral version of the intervention and then retesting them immediately after completion and 4 weeks later. Data collection took place in the Neurosciences Building, Warneford Hospital.

2. My Role in the Study

The study was designed by Dr. Michael Browning and Dr. Erdem Pulcu at the Department of Psychiatry at the University of Oxford, and aims to recruit at least 40 participants for each of three groups over about two years. This thesis is based on an interim analysis of this long term project which I worked on collecting data for seven months (until May 2017), and specifically analyses data from two groups: the “previously depressed” and control groups (*NB*: By May 2017, only 5 people in the currently depressed group had completed the study). I managed the recruitment of participants, collecting and analysing data as well as writing up this thesis. The period of subjects’ participation that this thesis included is from 11th November 2016 to 24th May 2017.

3. Participants

English-speaking people aged between 18 and 65 were recruited for the study by advertisement through social media and notices in several colleges in the University of Oxford. The general inclusion and exclusion criteria for the study were:

***Table 1*. Inclusion and Exclusion Criteria of the Study**

Those who …
a. Are fluent in English
b. Are not receiving any treatment for depression (including medication or talking treatments), or other treatment which may affect your mood and thinking
c. Don't have a diagnosis of bipolar disorder
d. Don't use street drugs (such as cannabis, ecstasy)
e. Don't have a medical condition which stops the pupils in your eyes getting bigger or smaller (in the study we measure the size of your pupils)
f. Haven't previously done a study which includes the same tasks we are using

In addition, participants in the previously depressed group had to have experienced at least two previous episodes of MDD, as confirmed using the SCID-V and participants in the control group had to have no current or previous diagnosis using the same measure. For participants in the previously depressed group, comorbid anxiety disorders were allowed.

4. Procedure

Participants visited the study centre a total four times (*Figure 3*). The 1st visit was for screening during which the researcher explained the details of the study,

obtained informed consent and confirmed the inclusion and exclusion criteria. During this visit the SCID-V was used to confirm participant group membership. The 2nd to 4th visits (Baseline Visit, Post-IBLT Visit & Follow-Up Visit) were testing sessions, during which participants were asked to complete several questionnaires and computerised cognitive tasks. The procedures are described in more detail in *Figure 3* and text below and appendices.

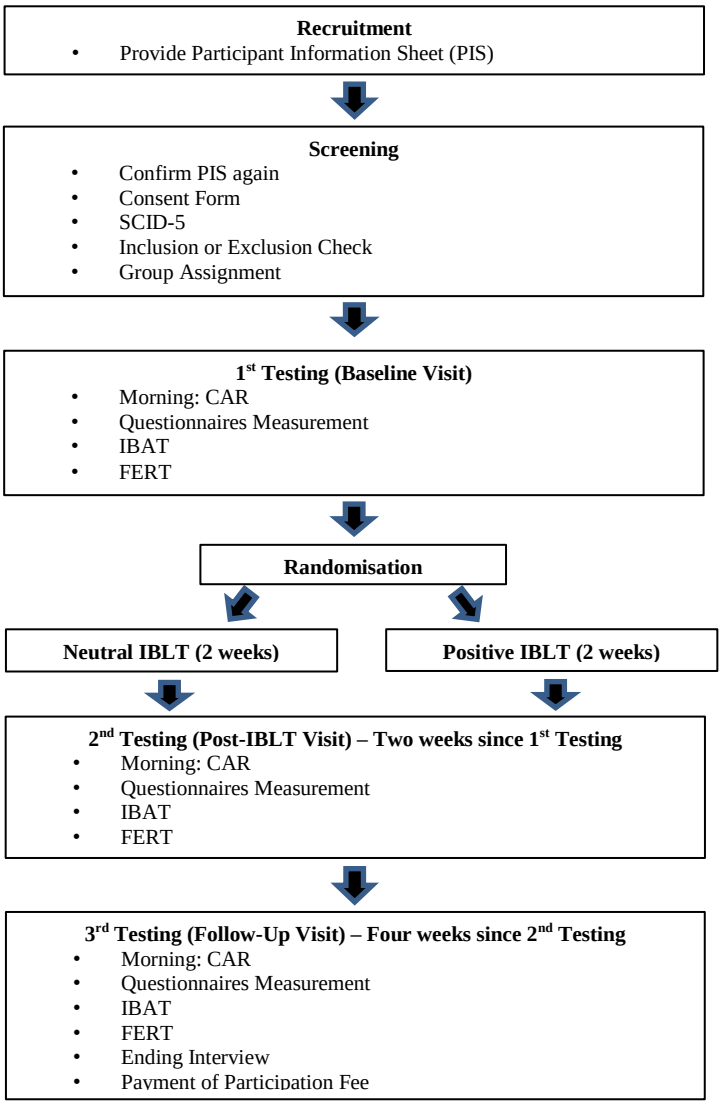


***Figure 3*. Flow Chart of the Study**

SCID-5, Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders; CAR, Cortisol Awaking Response; IBAT, Information Bias Assessment Task; FERT, Facial Expression Recognition Task; IBLT, Information Bias Learning Task.

Visit 1 (Screening): Researchers explained the detail of the study to participants who provided written informed consent. Inclusion and exclusion criteria were checked and the SCID-5 administered to check group membership. Participants were shown example versions of the learning tasks, and provided with saliva packets to use on the morning of next visit (for assessment of the cortisol awakening response).

Visit 2 (1st Testing; Baseline): Participants collected five saliva tubes after waking in the morning of the testing day (*[2]). Participants were interviewed about their mood using the Montgomery–Åsberg Depression Rating Scale (MADRS), and then asked to complete the 16-item Quick Inventory of Depressive Symtomatology (QIDS) (Rush et al., 2000), State-Trait Anxiety Inventory (STAI), Snaith–Hamilton Pleasure Scale (SHAPS) (Snaith, Hamilton, Morley, Humayan, Hargreaves & Trigwell, 1995), Depressive Attributions Questionnaire (DAQ) (Kleim, Gonzalo & Ehlers, 2011), and Ruminative Response Scale (RRS) (Nolen-Hoeksema & Morrow, 1991) questionnaires. Three computer based tasks were also completed, the Facial Expression Recognition Task (FERT), and two versions of the Information Bias Assessment Task (IBAT). These measures are described in more detail below. At the end, researchers emphasised to participants the importance of completing the training task (called the Information Bias Learning Task, IBLT) every day at home.

Between Visit 2 & 3: Participants were assigned randomly into either a neutral or positive IBLT group which determined the form of IBLT training completed (see

[2] The saliva samples will be analysed at the end of the study (about two-year period), therefore, I do not have access to the cortisol results for this thesis.

below for more detail). Participants were asked to complete this in their home every day between Visit 2 (1st Testing; Baseline Visit) and Visit 3 (2nd Testing; Post-IBLT Visit), which were about two weeks apart. If a participant missed the training on any day, the training system sent a notification e-mail to both the participant and researchers the next day. Researchers then contacted the participant to encourage them to complete the missed sessions.

Visit 3 (2nd Testing; Post-IBLT): This visit was identical to Visit 2 above. Participants did not complete the IBLT after Visit 3. In other words, they do nothing for the study between Visit 3 and 4.

Visit 4 (3rd Testing; Follow-Up): This visit was identical to Visit 2 above. After participants finished this visit, the researchers debriefed participants about the study and the payment form was processed. Participants were paid 100 pounds as basic participation fee, plus gains that they achieved during assessment tasks during three testing visit.

5. Questionnaire Measures

The study utilised questionnaires, described below, for assessing participants' mood and depressive status as well as associated anhedonia and cognitive symptoms. The symptoms of depression were measured using a self-reported form: QIDS-SR_{16} (Rush et al., 2003) questionnaire scores as well as a clinician rated tool, the MADRS questionnaire.

Montgomery–Åsberg Depression Rating Scale (MADRS): The MADRS is a 10-

item clinical questionnaire to rate severity of depression. The items asks the patient/participant's the past week's experience in questions of apparent sadness, reported sadness, inner tension, reduced sleep, reduced appetite, concentration difficulties, lassitude, inability to feel, pessimistic thoughts, and suicidal thoughts, and the interviewer rates defined scale (0, 2, 4 & 6 points) or between each of them (1, 3 & 5 points). This tool was first designed in 1979 (Montgomery & Åsberg, 1979).

Quick Inventory of Depressive Symptomatology (QIDS): The 16-item QIDS (Rush et al., 2000) is a shortened format of the 30-item Inventory of Depressive Symptomatology (IDS) (Rush et al., 1996), which assesses depressive symptoms and their severity. Both the IDS and QIDS have clinician (IDS-C_{30} & QIDS-C_{16}) and self-rating scales (IDS-SR_{30} & QIDS-SR_{16}) (Rush et al., 2003, p. 574), and the two versions of QIDS contain selected items of IDS that evaluate diagnostic symptoms of the American Psychiatry Association's Diagnostic and Statistical Manual of Mental Disorders - 4th edition (DSM-IV; APA, 1994) (Rush et al., 2003, p. 574). This study used self-rated version of QIDS (QIDS-SR_{16}), whose score range is between 0 and 27, rating symptoms of the past seven days. This consists of 16 items in nine domains: sad mood, concentration, self-criticism, suicidal ideation, interest, energy/fatigue, sleep disturbance, appetite/weight, and psychomotor agitation/retardation (Rush et al., 2003, p. 574).

Snaith–Hamilton Pleasure Scale (SHAPS): The SHAPS is an instrument to assess anhedonia (the inability to experience pleasure), constructed by Snaith and researchers in 1995 (Snaith et al., 1995). This tool is based upon an idea that low hedonic tone plays an important role in mood disorders (Snaith et al., 1995, p. 99).

The questionnaire consists of 14 items asking the person to agree or disagree with descriptions of pleasurable experience in recent days, with four responses “strongly disagree”, “disagree”, “agree” and “strongly agree”, which covers four kinds of experience in pleasure interest/pastimes, social interaction, sensory experience and food/drink (Snaith et al., 1995, p. 102-103).

State-Trait Anxiety Inventory (STAI): The State-Trait Anxiety Inventory (STAI) consists of two 40-item questionnaires that measures two aspects of anxiety. One is anxiety in the temporary and episodic form, which is transient (state anxiety), and the other one is anxiety as a more stable trait (trait anxiety), which measures an individual’s general tendency to experience symptoms of anxiety (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983). This self-rating tool gives 20 items to each of the State Anxiety Scale and the Trait Anxiety Scale: The State Anxiety Scale asks how the person feels at the moment in terms of apprehension, tension, nervousness, worry, and activation/arousal of the automatic nervous system, with four choices of “not at all”, “somewhat”, “moderately so” and “very much so”, whereas the Trait Anxiety Scale evaluates the respondent’s general state in calmness, confidence, and security, with four options of “almost never”, “sometimes”, “often” and “almost always” (Julian, 2011, p. S467). The STAI was first introduced in 1970 (Spielberger, Gorsuch & Lushene, 1970) (the STAI Form X), and then developed in 1983 (Spielberger et al., 1983) as the STAI Form Y (Julian, 2011, p. S467). In this thesis, only the trait-STAI data was analysed.

Depressive Attributions Questionnaire (DAQ): The DAQ is a measure of depressogenic attributions (Kleim, Gonzalo & Ehlers, 2011, p. 381). This includes 16 items assessing how the respondent thinks or feels about negative and positive

events, with a scale ranging from 0 (“not at all”) to 4 (“very strongly”) (Kleim, Gonzalo & Ehlers, 2011, p. 376, 383).

Ruminative Response Scale (RRS): The RRS is one of the subscales of the Response Styles Questionnaire (Nolen-Hoeksema & Morrow, 1991). The term “rumination” indicates a tendency of “responding to distress that involves repetitively and passively focusing on symptoms of distress and on the possible causes and consequences of these symptoms (Nolen-Hoeksema, Wisco & Lyubomirsky, 2008, p. 400)", and rumination has been studied for its relationship with depression and negative cognitive styles (Nolen-Hoeksema, Parker & Larson, 1994, p. 93-94, 96-100; Spasojević & Alloy, 2011, p. 26, 30-33). The RRS is made up of 22 items asking about different aspects of rumination (Treynor, Gonzalez & Nolen-Hoeksema, 2003, p. 248). Higher levels of rumination have been associated with PTSD (Nolen-Hoeksema & Morrow, 1991), worry (Segerstrom, Tsao, Alden & Craske, 2000), anxiety (Nolen-Hoeksema, 2000; Cox, Enns & Taylor, 2001) and depression (Nolen-Hoeksema et al., 1994; Treynor, Gonzalez & Nolen-Hoeksema, 2003).

6. Facial Expression Recognition Task (FERT)

The study adopted a facial expression recognition task, called the “face emotion recognition task” (FERT) (Harmer et al., 2003a; Harmer et al., 2003b; Harmer et al., 2009). The task assesses how participants classify the emotion that is displayed on the screen of face stimuli of six emotions: anger, disgust, fear, happy, sad, surprise (Ekman and Friesen, 1971). The facial images were morphed with

different intensity levels in steps of 10% between 0% (neutral) and 100% (each emotional standard). This morphing procedure increases the ambiguity of the emotions and the sensitivity of the task when detecting between subject differences. Four samples for each emotion with ten intensities (6 emotions × 10 intensities × 4 examples = 240 stimuli), and ten neutral stimuli were given to participants (total 250 stimuli). The task presented the stimuli in a random order for a brief duration (500ms) to participants, and the participant was asked to respond to which kind of emotion it was by choosing one of seven options (anger, disgust, fear, happy, sad, surprise & neutral). In previous studies, antidepressants have been found to alter performance of this task (Harmer et al., 2004). For example, citalopram improved correct identification of fear (Harmer et al., 2003a), and both citalopram and reboxetine did so to happy stimuli (Harmer et al., 2003b). The participants' performance in this task was measured as 'accuracy' (proportion of faces showing an emotion which were correctly labelled), 'misclassification' (proportion of faces not showing an emotion which were incorrectly labelled as showing it) and 'reaction time' (mean reaction time for correctly labelled faces). As the specific prediction associated with this study was that previously depressed participants would show a negative bias and that this would be ameliorated by the IBLT training, the FERT outcomes were simplified by calculating the mean of the positive (happy and surprise) and the negative (sad, anger, fear, disgust) emotions. This allowed an analysis of "positive relative to negative" emotions which produces a direct estimate of affective bias. This bias analysis was complemented by an analysis of the separate emotions which is able to assess for the presence of emotional specificity in any more general effect.

7. Information Bias Learning Task (IBLT) & Information Bias Assessment Task (IBAT)

The current study utilised a task, the Information Bias Assessment Task (IBAT), which assessed how participants' learned from positive and negative outcomes and how they adapted their learning to the information content of those outcomes. Participants completed this task during each testing visit (Visit 2, Visit 3 & Visit 4). A second version of this task, designed to induce a bias in learning rather than simply measure it, the information bias learning task (IBLT), was completed by participants every day at home during the period between the first two testing visits (Visit 2 & Visit 3).

Common Principle: Participants are presented with two different shapes on a screen, and have to choose one of these two. On every trial, following the participant's choice a "win" and a "loss" outcome appears behind one of the shapes. If the participant has chosen the shape associated with the "win" outcome they are given 15p, if they have chosen the shape associated with the "loss" outcome then they lose 15p. The two outcomes are independent which means that they may appear behind the same or different shapes. This also means that knowing where one is (e.g. the win) will not tell you where the other (e.g. loss) is. Participants' are asked to learn from the outcome of previous trials which shape is better for them, in terms of earning the money or points. In other words, they have to learn where they think the win will appear and where they think the loss will appear. They then have to use these estimates to guide their choice of shape.

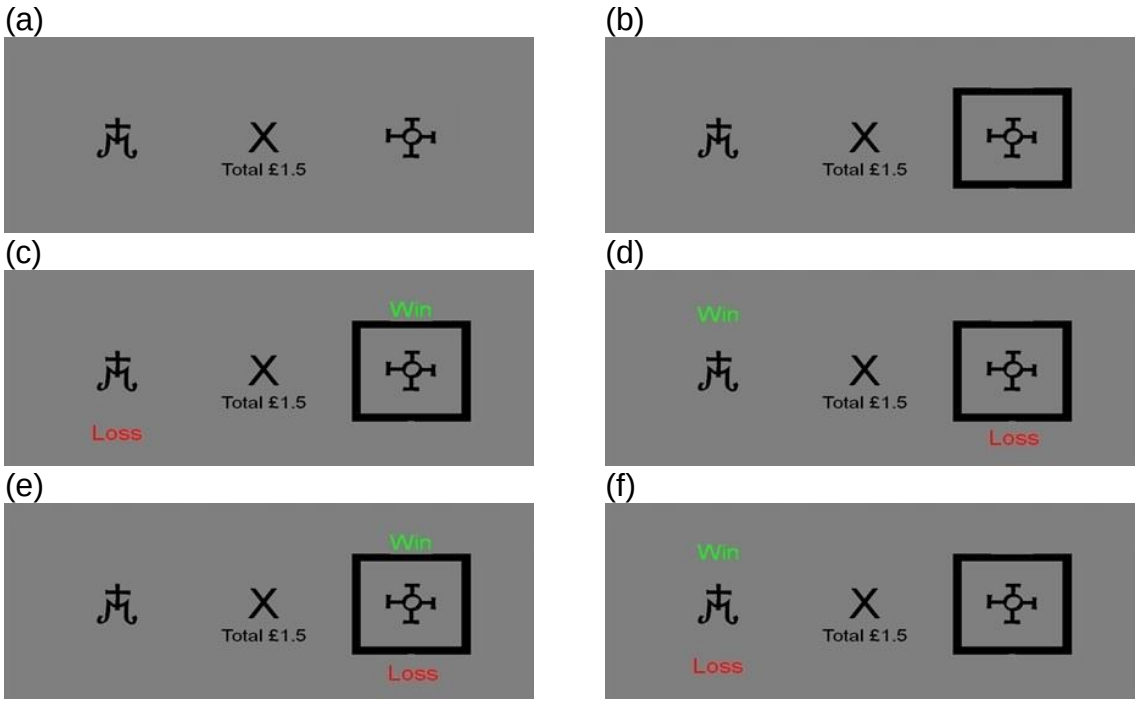


***Figure 4*. One Trial and Possible Outcomes of the Information Bias Tasks (Both IBAT & IBLT)**

(a): The participant was given two different shapes on left and right side at the beginning of each trial. (b): Then a participant chose one of two shapes. The image shows the example that the participant chose the right one. (c), (d), (e), (f): Once the participant chose the shape, the outcomes of "win" and "loss" appeared shortly, and the four images show possible outcomes for a given trial.

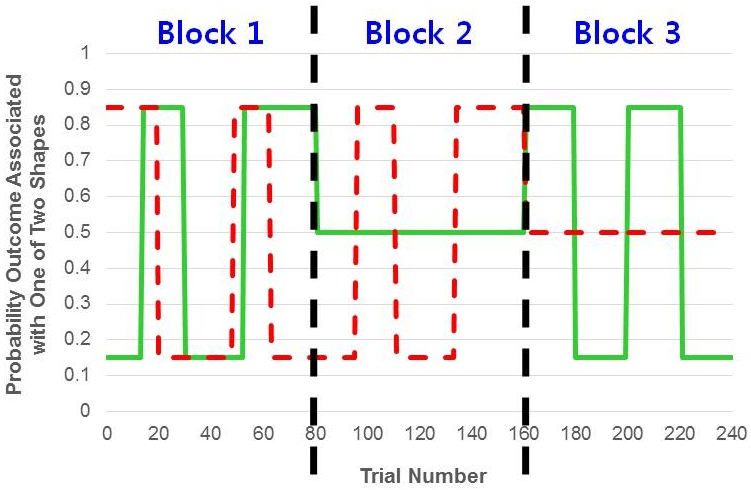


***Figure 5*. Structure of the Information Bias Assessment Task (IBAT)**

The IBAT consists of 240 trials split into three blocks. The basic information of the figure is the identical with the one in *Figure 2* above: the x-axis indicates the number of trials, whereas the y-axis is the probability (*p*) of outcomes "win" and "loss" for a certain shape. The probability of two outcomes for the other shape is "1-*p*". In the 1st block, probabilities for two outcomes were volatile. In the 2nd block, only the outcome probability for "loss" was volatile, but the one for "win" was stable. This was the opposite in the 3rd block. For the training task (IBLT), participants completed one block repeatedly at home over two weeks. For the positive IBLT, participants completed the 3rd block above. For the neutral IBLT, they completed the 1st block.

IBAT: This is completed in each testing session, and participants have the final total sum they earn during the task added to their participation fee. The task consists of a total of 240 trials split into three separate blocks (*Figure 5* above). The volatility of the two outcomes is manipulated between the blocks: In the first block both positive and negative outcomes are volatile, in the second block the positive outcome is volatile, but the negative outcome is stable, and in the third block the negative outcome is volatile, but positive outcome is stable. A rest session occurs between the blocks and the stimuli used in the task differs for the blocks. All three blocks are completed by participants during each test session, but the order of the last two blocks is counterbalanced across participant and test session. The choices made by participants during this task is analysed using a computational model which is described later.

IBLT: During the two week period between the first two test sessions, participants were asked to complete the IBLT every day in their home. The structure of the IBLT is the same as the IBAT described above, except that reward and loss outcomes lead to participants winning or losing points, not money. The critical manipulation in the IBLT is that participants' randomised group determined the schedule of the task. Participants in the positive IBLT group completed a version of the task in which win outcomes were volatile and losses stable (i.e. Block 3 in *Figure 5* above), whereas participants in the neutral IBLT group completed a version of the task in which both wins and losses were volatile (i.e. Block 1 from *Figure 5* above). The rationale for this task is that participants in the positive group have to repeatedly complete a task in which they must learn more from positive than negative outcomes, that is they must develop a positive learning bias,

whereas the neutral group must treat positive and negative outcomes equivalently. In this study we assessed the degree to which this bias training succeeded in altering learning style and the extent to which this effect generalised to other measures of cognition and mood.

8. Computational Model

The current study is based on a theory of reinforcement learning which underlines how behaviour adapts with experience to maximise rewards and minimise punishments (Sutton & Barto, 1998). The reinforcement model used in this study has two types of parameter: (1) the “learning rate”, which describes how rapidly subjects alter their value estimates with prediction errors, and (2) the “temperature”, which describes the degree which subjects’ decisions are guided by their estimates of the stimulus values (Raymond, Steele & Seriès, 2017. p. 2).

The main focus of this study is the assessment of how participants learn from positive relative to negative outcomes. To achieve this, the study estimated the learning rate and temperature parameters separately for the two outcomes (“win” & “loss”), in each of the three blocks. The model used is based on that described in Behrens et al. (2007) and Browning et al. (2015).

First, the model estimated the probability that the “win” and “loss” outcomes would be associated with a specific shape (“Shape A”), using two Rescorla-Wagner learning rules (Rescorla & Wagner, 1972; Pulcu & Browning, 2017):

***Equation 1*. Equations for Estimating the Probability that Each Outcome would be Associated with "Shape A" (Rescorla & Wagner, 1972; Pulcu & Browning, 2017)**

$$rwin_{(i+1)} = rwin_{(i)} + \alpha win \times (winout_{(i)} - rwin_{(i)})$$
$$rloss_{(i+1)} = rloss_{(i)} + \alpha loss \times (lossout_{(i)} - rloss_{(i)})$$

$rwin_{(i)}$: Estimated probability that the "win" is associated with shape A on trial i

$winout_{(i)}$: A coded variable for whether the "win" was associated with shape A (a value of 1) or shape B (a value of 0)

αwin: The learning rate for "wins", a free parameter in the model

$rloss$, $lossout$ & $\alpha loss$ are the same variables for the "loss" outcomes.

The equations above describes how the estimated probability that each outcome ("win" & "loss") is associated with shape A which is updated following the outcome of each trial. The estimate for shape B is simply 1- that for shape A. The probability $rwin_{(i)}$ is initially set as 0.5. During learning this estimate is added to the product of $rwin_{(i)}$ (the learning rate) and the difference between $winout_{(i)}$ and $rwin_{(i)}$ (the prediction error). Note this leads to a large updating of belief if either what a participant expected ($rwin_{(i)}$) is quite different from what he/she actually observes ($winout_{(i)}$), and/or the learning rate is large. An identical process is used to estimate the probability that the loss will be associated with shape A.

These updated values were then combined in a softmax function which calculates the probability that the model will chose shape A (see *Equation 2* below). The softmax function had two additional free parameters, the inverse temperatures for wins (βwin) and losses ($\beta loss$). When these parameters have a large value, the model's choice is driven to a greater extent by the learned stimuli values, whereas when the values are low the choice is more random.

***Equation 2*. Choice Probability Equation (Pulcu & Browning, 2017)**

$$PchoiceA_{(i)} = \frac{1}{1 + exp^{-(\beta win \times rwin_{(i)} - \beta loss \times loss_{(i)})}}$$

$PchoiceA_{(i)}$: Estimated probability that a participant choose a shape A on trial i

βwin & $\beta loss$: Estimated inverse temperature for "win" & "loss", respectively

Overall, the model has four free parameters (separate learning rate and inverse temperature parameters for win and loss outcomes). The values of these parameters were estimated from participant choice separately for each participant and each task block. Parameter estimation was achieved by calculating the full joint probability of the parameters given the participant's choice. The distribution of the four separate parameters was then calculated by marginalization and the expected value of each parameter derived from this distribution (Behrens et al., 2007; Browning et al., 2015; Pulcu & Browning, 2017). Note that in order to estimate the parameters across the continuous real line (i.e. from plus to minus infinity) an inverse logit transform was used for learning rate and a log transform was used for the inverse temperature (Pulcu & Browning, 2017).

9. ANOVA & T-tests Analysis

The current study used analysis of variance (ANOVA) to test for differences between groups and baseline as well as the effect of the IBLT training over time. Baseline group differences were tested using either univariate ANOVA (for questionnaire measures) or repeated measures ANOVAs (for IBAT and FERT tasks). In these analyses participant group was included as a between subject

factor. To assess the effect of the training intervention, IBLT, repeated measure ANOVAs were used with between subject factors of “subject group” and “training group (IBLT Type)”, and within subject factor of “time (testing visit)”. For the IBAT and FERT task additional task related within subject factors were included as above. Significant or marginal effects were followed up using standard post-hoc analyses. The justification for including marginal effects in the analysis was that the current thesis describes an exploratory analysis of interim study data. Where assumptions of sphericity were not met Greenhouse-Geisser corrections were used to adjust the degrees of freedom of the ANOVAs.

Ⅲ. Result

1. Sample Demographics

A total of 34 participants had been recruited to the study by May 2017, which was the cut off for inclusion in this thesis: 16 and 18 were recruited to the previously depressed group, and control group respectively. Some participants had not completed all three study test visits by this cut off and so had incomplete data: *Previously Depressed Group*: One participant had not completed either the last or second last visits (Post-IBLT Visit & Follow-Up Visit), and another two participants had not completed the last visit. *Control Group*: Five participants had completed the last visit. Groups were well matched for gender, age and the number of IBLT sessions completed (see *Table 2* below). Of note the IBLT groups (positive & neutral) significantly differed in the average points that participants obtained during IBLT training: The two groups completing neutral IBLT achieved a mean of 35.05 and 32.79 points, while the groups completing the positive IBLT achieved a mean of 25.53 and 24.29 points. This difference results from the fact that in the neutral IBLT participants could learn from both positive and negative outcomes, whereas in the positive IBLT participants could only learn from positive outcomes.

***Table 2*. Baseline Characteristics**

	Previously Depressed Group with Neutral IBLT	Previously Depressed Group with Positive IBLT	Control Group with Neutral IBLT	Control Group with Positive IBLT	Univariate ANOVA (*p*)		
					Difference between Subject Groups	Difference between IBLT Types	Interaction of Subject Group & IBLT Type
Number	8	8	9	9	…	…	…
Gender	7 female	6 female	7 female	6 female	0.559	0.446	0.964
Age (Years)	31.88 (12.11)	31.63 (11.07)	30.44 (10.94)	29.78 (8.71)	0.660	0.902	0.955
Number of Previous Episodes of Depression	2.43 (0.79) (*[3])	2.75 (0.89)	0 (0)	0 (0)	…	0.318	…
Number of IBLT Sessions Completed	14.50 (1.69)	14.75 (2.19)	15.89 (1.45)	15.00 (0.87)	0.145	0.564	0.307
Mean of Daily Points During IBLT	35.05 (3.30)	25.53 (2.11)	32.79 (5.46)	24.29 (5.21)	0.248	0.000	0.734
QIDS (Baseline)	4.00 (3.16)	3.50 (1.69)	3.67 (3.00)	2.44 (2.96)	0.473	0.375	0.708
SHAPS (Baseline)	52.63 (4.41)	46.50 (3.12)	47.89 (7.39)	46.89 (8.49)	0.329	0.114	0.251
Trait-STAI (Baseline)	34.88 (8.18)	45.00 (6.52)	29.89 (7.10)	28.89 (7.62)	0.000	0.082	0.036
DAQ (Baseline)	11.63 (8.77)	21.50 (6.63)	10.22 (7.12)	9.56 (6.65)	0.013	0.077	0.045
RSS (Baseline)	46.38 (14.46)	54.38 (9.53)	34.78 (6.55)	35.78 (13.62)	0.001	0.261	0.380
MADRS (Baseline)	4.75 (2.37)	4.63 (0.80)	3.78 (1.77)	3.33 (1.86)	0.535	0.875	0.930

Data are given as mean (standard deviation) unless indicated. ANOVA, Analysis of Variance; IBLT, Information Bias Learning Task; SEM, Standard Error of the Mean

[3] One extreme outlier number (60 episodes) reported by the participant No. 2011 was excluded.

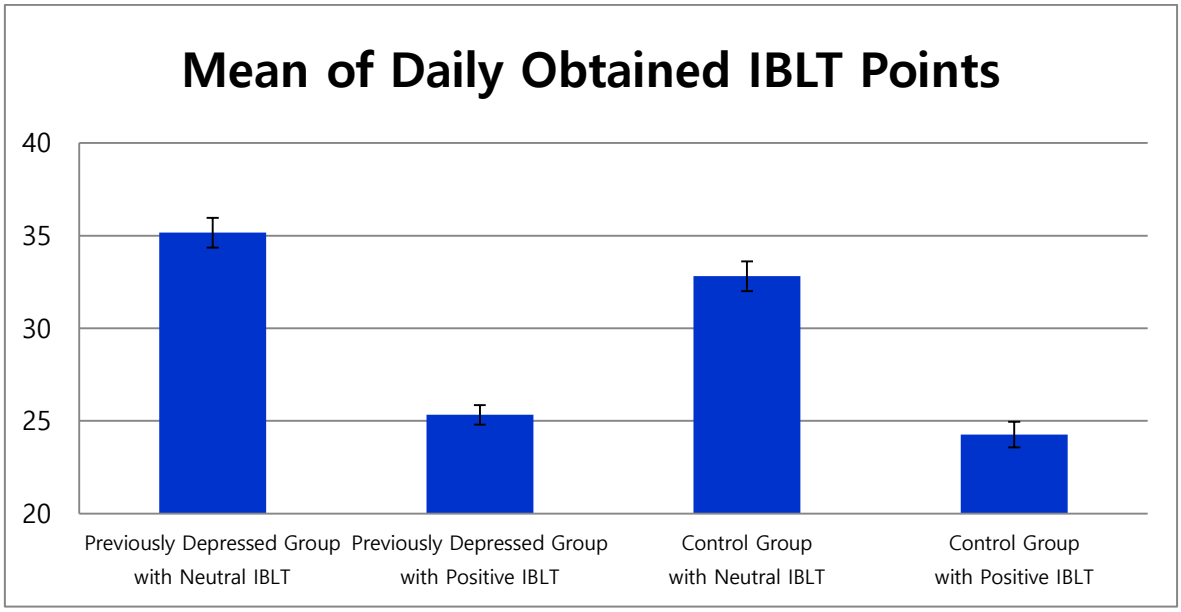


***Figure 6*. Average Points of Daily Won During IBLT by 4 Groups**

Mean (SEM) of IBLT points during about 2 weeks home training separately for 4 groups

2. IBAT Task: Learning Rate

The learning rates of participants in the IBAT task was analysed using repeated measures ANOVAs with Block (Block 1= both outcomes volatile, 2= losses volatile & 3=wins volatile) and Outcome ("win" & "loss") as within subject factors. For the baseline analysis subject group was included as a between subject factor. For analysis of the effect of IBLT training, IBLT group was included as an additional between subject factor, and visit (1, 2 & 3) as an additional within subject factor.

2.1. Baseline Comparison

A repeated measures ANOVA with Greenhouse-Geisser correction found a significant effect of Block × Outcome (*Figure 7*; $F(1.42, 45.30)=16.52$, $p<0.001$)

which reflects the fact that all participants adjusted their learning rates as expected—using a higher learning rate for more informative outcomes. However, there was no significant effect of group, with main effect of Subject Group ($F(1)=0.39$, $p=0.54$) as well as the Block × Outcome × Subject Group ($F(1.42, 42.63)=0.12$, $p=0.82$), Block × Subject Group ($F(2, 60)=0.35$, $p=0.70$), and Outcome × Subject Group ($F(1, 30)=0.002$, $p=0.96$) being non-significant. Thus there was no evidence for differential learning during the IBAT task in the 2 subject groups.

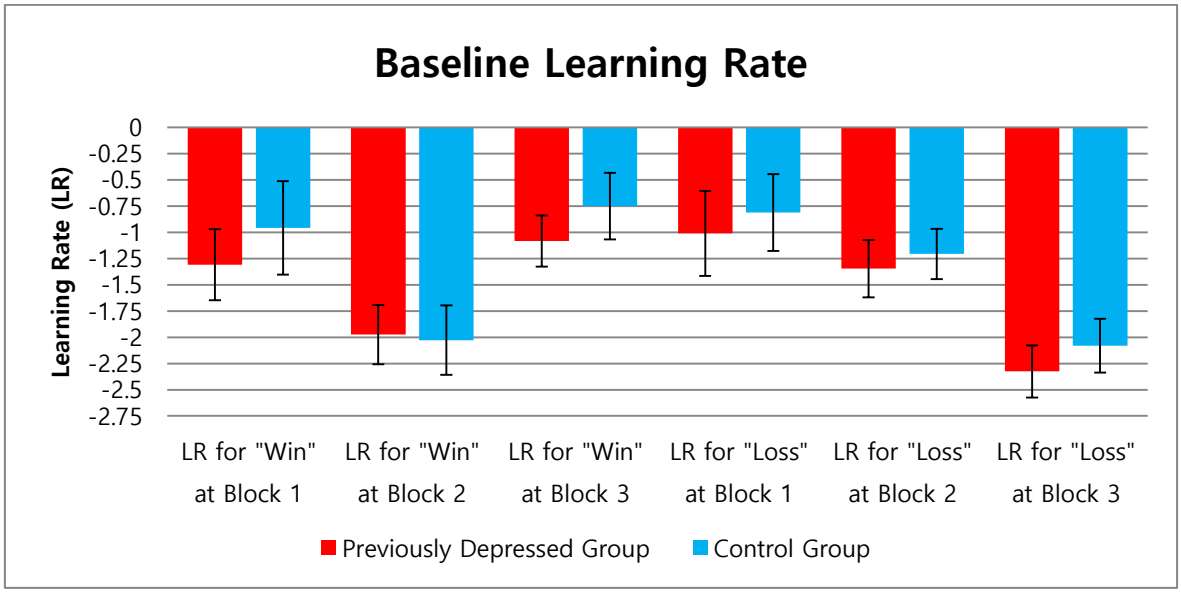


***Figure 7*. Learning Rates (LR) of the Two Participant Groups at Baseline**

Mean (SEM) of learning rates for previously depressed (red) and control (blue) participants. The first three pairs of columns report the learning rates for positive outcomes ("wins"), with the last three reporting the learning rates for negative outcomes ("losses"). Note the learning rates are in inverse logit space and therefore do not lie between 0 and 1.

2.2. IBAT Task: The Influence of IBLT over Time on Learning Rate

In order to test whether learning training (IBLT) affected learning rates over time, a

repeated measures ANOVA was conducted. The Visit × Outcome × IBLT Type interaction showed a trend level effect, $F(2, 50)=2.98$, $p=0.06$ which was not modified by subject group (Visit × Outcome × Subject Group × IBLT Type, $F(2, 50)=1.94$, $p=0.15$), or block type (Visit × Outcome × Block × IBLT Type, with Greenhouse-Geisser correction, $F(2.62, 65.52)=0.82$, $p=0.47$) however. The interaction of Visit × Outcome × Block × Subject Group × IBLT Type was not significant, with Greenhouse-Geisser correction, $F(2.62, 65.52)=0.94$, $p=0.42$. In other words, subject group and block type did not influence the effect of the IBLT training, although there was a marginal effect of the IBLT training itself. For further analysis, post hoc tests of the Visit x Outcome x IBLT Type were conducted.

In order to assess the effect of IBLT Type (neutral & positive IBLTs), repeated measures ANOVAs were conducted separately for the two different outcomes (“win” & “loss”), between baseline and post-IBLT visits, and between post-IBLT and follow-up visits.

- The Effect of IBLT upon Learning Rate for “Win” Between Baseline & Post-IBLT Visits

A repeated measures ANOVA reported that the effect of Visit was significant, $F(1, 30)=14.50$, $p=0.001$, but the Visit × IBLT Type was not significant, $F(1, 30)=0.13$, $p=0.72$. The main effects of “Subject Group”, “IBLT Type”, and “Subject Group × IBLT Type” were not significant ($p>0.17$). As can be seen from *Figure 8*, positive learning rate increased across the training period in both groups.

- The Effect of IBLT upon Learning Rate for “Loss” Between Baseline & Post-IBLT Visits

Two trend level effects were found by a repeated measures ANOVA. The Visit Effect was marginal, $F(1, 30)=3.87$, $p=0.06$, and the interaction of Visit × IBLT Type was marginal, $F(1, 30)=3.28$, $p=0.08$.

Analysing the change in loss learning rate over time, separately for participants in the neutral and positive IBLT conditions revealed a significant increase for neutral IBLT, $F(1, 15)=11.44$, $p=0.004$ with no significant change for positive IBLT, $F(1, 15)=0.01$, $p=0.93$. Lastly, analysing the difference between the two IBLT groups separately for the two study visits found not significant difference for either ($p>0.17$). In summary, in the neutral IBLT condition both win and loss learning rate increased across training, whereas for the positive IBLT condition only the win learning rate increased. This reflects the intended effect of the training in that positive IBLT included volatile (informative) wins and stable (uninformative) losses, where both outcomes were volatile during neutral IBLT.

- The Effect of IBLT upon Learning Rate for “Win” Between Post-IBLT & Follow-Up Visits

A repeated measures ANOVA reported that the Visit effect was not significant, $F(1, 25)=0.24$, $p=0.63$, but the Visit × IBLT Type effect was significant, $F(1, 25)=4.56$, $p=0.04$.

As can be seen in *Figure 8* there was a marginal increase in win learning rate in the neutral IBLT condition, $F(1, 13)=3.74$, $p=0.08$, with no significant change for positive IBLT $F(1, 12)=1.25$, $p=0.29$.

- The Effect of IBLT upon Learning Rate for “Loss” Between Post-IBLT & Follow-Up Visits

No significant effects of “Visit” or “Visit × IBLT Type” were found for loss learning rates in the follow up period ($p>0.23$).

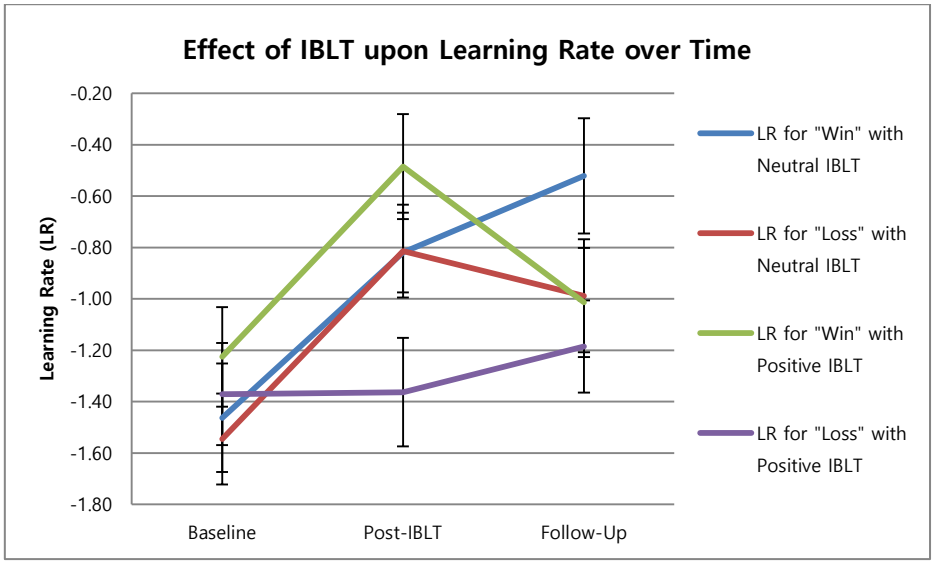


***Figure 8*. The Effect of IBLT upon Learning Rate over Time**

Mean (SEM) of the learning rates are shown for each IBLT group and the each outcome (wins and losses). Note that learning rate is shown in inverse logit space and so is not limited to 0 to 1.

- The Effect of IBLT upon Learning Rate: Summary

To sum up, the effect of IBLT upon learning rate across training was (at a trend level) that expected from the design of the training task. These effects began to erode across the follow-up period.

3. IBAT Task: Inverse Temperature

Analyses of the inverse temperatures of the IBAT used the same strategy as that for the learning rates.

3.1. Baseline Comparison

A repeated measures ANOVA did not find any significant effect (p>0.25): None of the Block × Subject Group effect, $F(2, 64)=1.42$, $p=0.25$, the Block × Outcome effect, with Greenhouse-Geisser correction, $F(1.53, 48.95)=0.88$, $p=0.40$, the Outcome × Subject Group effect, $F(1, 32)=1.01$, $p=0.32$, and the Block × Outcome × Subject Group effect, with Greenhouse-Geisser correction, $F(1.53, 48.95)=0.21$, $p=0.75$, were significant. The main effect of subject group was not significant, $F(1, 1.44)$, $p=0.24$, too.

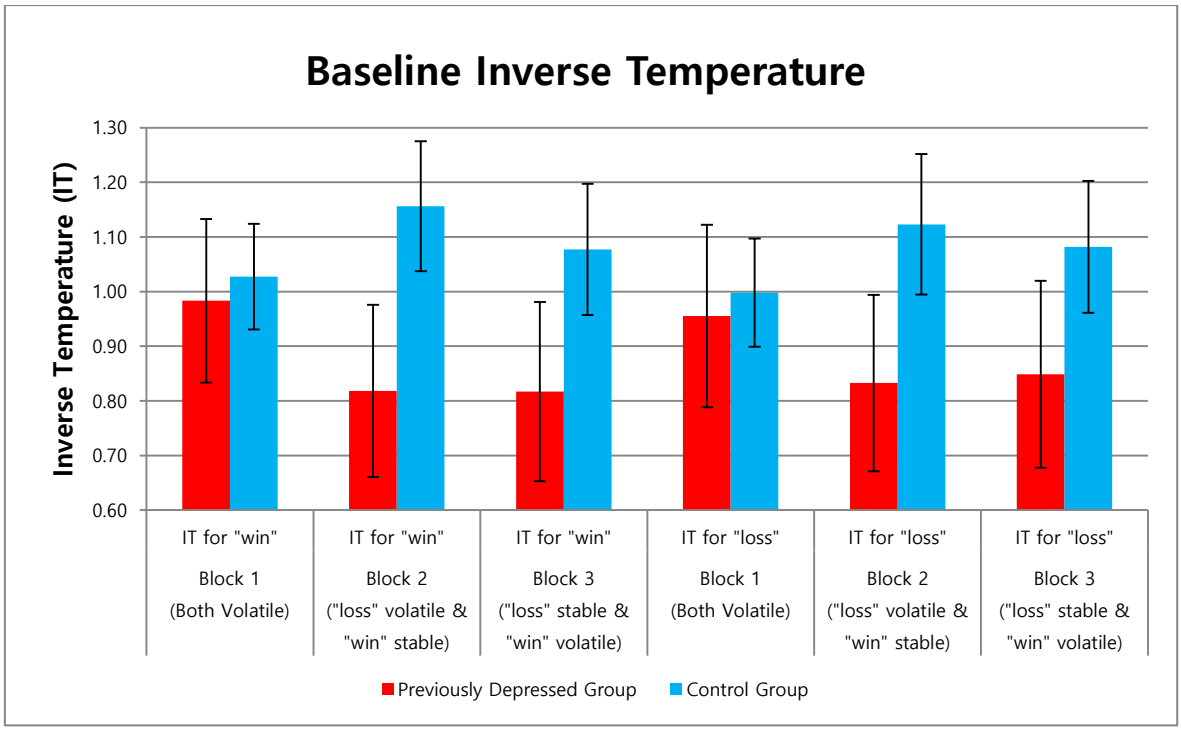


***Figure 9*. Inverse Temperatures (IT) of the Two Participant Groups at Baseline**

Mean (SEM) of inverse temperatures for previously depressed (red) and control (blue) participants. The first three pairs of columns report the learning rates for positive outcomes (“wins”), with the last three reporting the inverse temperatures for negative outcomes (“losses”).

3.2. IBAT Task: The Influence of IBLT over Time on Inverse Temperature

A repeated measures ANOVA revealed several significant and marginal findings. There was a significant effect of Visit × Block × IBLT Type, $F(4, 112)=3.50$, $p=0.01$, and a marginal effect of Visit × Outcome × IBLT Type, $F(2, 56)=3.05$, $p=0.06$. In addition, a marginal value of Visit × Block × Subject Group × IBLT Type, $F(4, 112)=2.12$, $p=0.08$, and another marginal effect of Visit × Block × Outcome × IBLT Type, with Greenhouse-Geisser correction, $F(2.68, 74.93)=2.40$, $p=0.08$, were reported.

Given the complex nature of these results, the analysis focuses on the interaction of IBLT type and visit. The highest level interactions found were Visit x Block x Subject Group x IBLT Type and Visit x Block x Outcome x IBLT Type. These interactions are described below by separately examining the visit x IBLT x block interactions for the four groups (i.e. control and previously depressed subjects who completed the positive and neutral IBLT tasks).

- The Effect of IBLT upon Inverse Temperature in the Previously Depressed Group

Overall the inverse temperature of both wins and losses in the previously depressed group (*Figures 10 and 11*) responded similarly. For positive IBLT there was a general increase in inverse temperature in all three blocks across training (*Figure 10 & Figure 11*), whereas for neutral IBLT temperatures (*Figure 12*) there was a general increase in inverse temperature in Blocks 1 and 3 (both volatile and win volatile) and a decrease in Block 2 (loss volatile). A detailed analysis of these

effects is presented below.

First, the change across the training period was tested with a repeated measures ANOVA of the inverse temperature for wins in the previously depressed group between the baseline and post-IBLT visits (*Figure 10*). This analysis found a marginal Visit × Block × IBLT Type interaction $F(2, 28)=2.89$, $p=0.07$, with the same analysis of loss inverse temperature (*Figure 10*) being significant, $F(2, 28)=4.15$, $p=0.03$. The same analysis was conducted again for “win” and “loss” outcomes separately in each of the three task blocks between baseline and post-IBLT visits: No significant effects of IBLT over time were found for either win ($p>0.27$) or loss ($p>0.12$) outcomes in the separate blocks.

Next the change between post-IBLT and follow-up visits was analysed. A significant, visit x block x IBLT interactions was found for both win $F(2, 26)=4.13$, $p=0.03$, and loss outcomes, $F(2, 26)=4.30$, $p=0.02$. Analysing the three blocks separately,: For “win” outcomes, the Visit × IBLT Type was marginal in Block 1, $F(1, 13)=4.62$, $p=0.05$, significant in Block 2, $F(1, 13)=8.89$, $p=0.01$, but not significant in Block 3, $F(1, 13)=0.80$, $p=0.38$. For “loss” outcomes, the Visit × IBLT Type was marginal in all three blocks. There were marginal effects in Block 1, $F(1, 13)=4.07$, $p=0.07$, in Block 2, $F(1, 13)=4.08$, $p=0.07$, and Block 3, $F(1, 13)=3.33$, $p=0.09$.

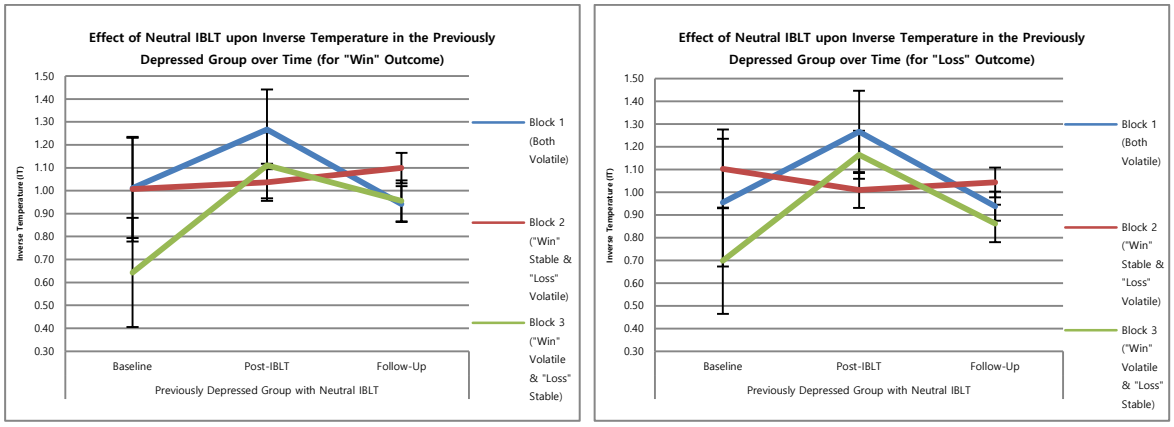


***Figure 10*. Effect of Neutral IBLT upon Inverse Temperature in the Previously Depressed Group for "Win" & "Loss" Outcomes over Time**

The mean (SEM) of the win inverse temperatures (in log space) are displayed for each visit. *Figure 10* (left) describes the effect of neutral IBLT on inverse temperature of previously depressed group, for "win" outcomes in three testing visits. *Figure 10* (right) displays the same result for "loss" outcome. As can be seen in these two figures, inverse temperatures in Block 1 & 3 for both outcomes increased across training and then reduced in the follow-up period. The difference in Block 2 between in two outcomes can be seen in the change between baseline & post-IBLT: figure for "win" outcome showed a slight increase while the one for "loss" showed a decrease.

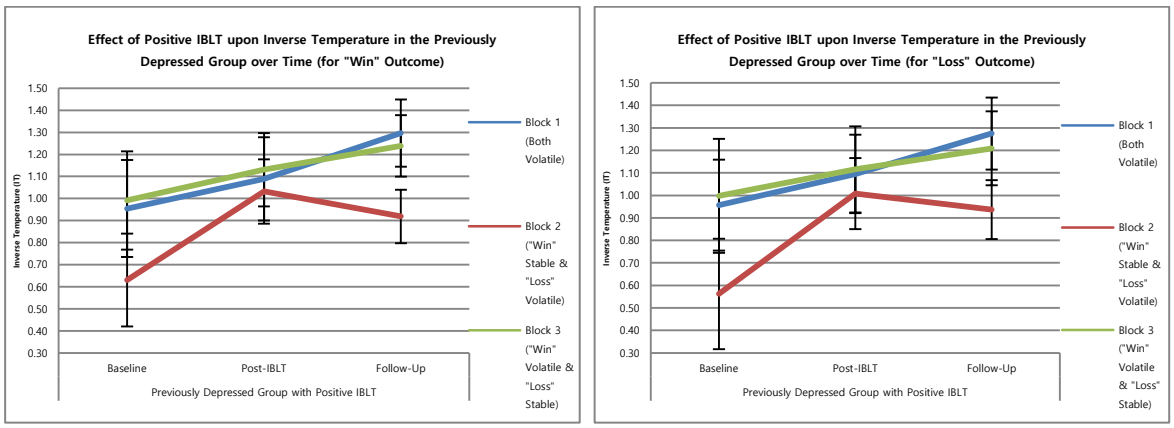


***Figure 11*. Effect of Positive IBLT upon Inverse Temperature in the Previously Depressed Group for "Win" & "Loss" Outcomes over Time**

As can be seen, "win" (left) & "loss" (right) outcomes with positive IBLT appear to have very similar effects on loss inverse temperatures across the study.

- The Effect of IBLT upon Inverse Temperature in the Control Group

Looking first across the training period: no significant Visit × Block × IBLT Type effect was found for either “win” $F(2, 32)=1.45$, $p=0.25$, or “loss” outcomes, $F(2, 32)=2.03$, $p=0.15$.

Next looking across the follow up period, no Visit x Block x IBLT Type interaction was found for “win” outcomes, $F(2, 30)=2.31$, $p=0.12$,with only a marginal effects for “loss”, $F(2, 30)=3.05$, $p=0.08$. Further analysis of this last effect revealed a significant Visit × IBLT Type effect for Block 1, $F(1, 15)=9.62$, $p=0.01$, a marginal effect in Block 2, $F(1, 15)=3.74$, $p=0.07$, but not effect in Block 3, $F(1, 15)=0.56$, $p=0.47$.

A complex set of block specific effects of IBLT training on inverse temperatures was found, with the effects being somewhat greater in the previously depressed than control group. Generally, in previously depressed patients, the inverse temperature increased across training for both outcomes in all Blocks in those completing positive IBLT, whereas for those completing neutral IBLT inverse temperatures for both wins and losses increased in Blocks 1 and 3 and decreased in Block 2. Unlike the effect on learning rate reported above it is difficult to understand this effect in terms of the design of the training task, although it may reflect a general tendency for participants who completed the neutral IBLT to become more engaged specifically in blocks in which the wins were informative.

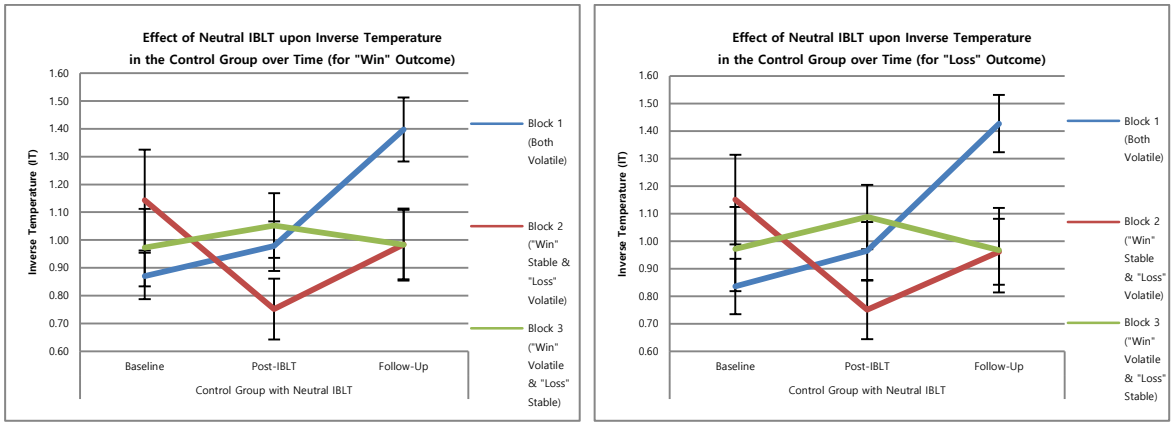


***Figure 12*. Effect of Neutral IBLT upon Inverse Temperature in the Control Group for "Win" & "Loss" Outcomes over Time**

Results show mean (SEM) of inverse temperature separately for the three task blocks (in log space). *Figure 12* (left) illustrates the effect of neutral IBLT on inverse temperature of control group for "win" outcome in three testing visits. *Figure 12* (right) is the same for "loss" outcome. Both figures showed changes over time in each block.

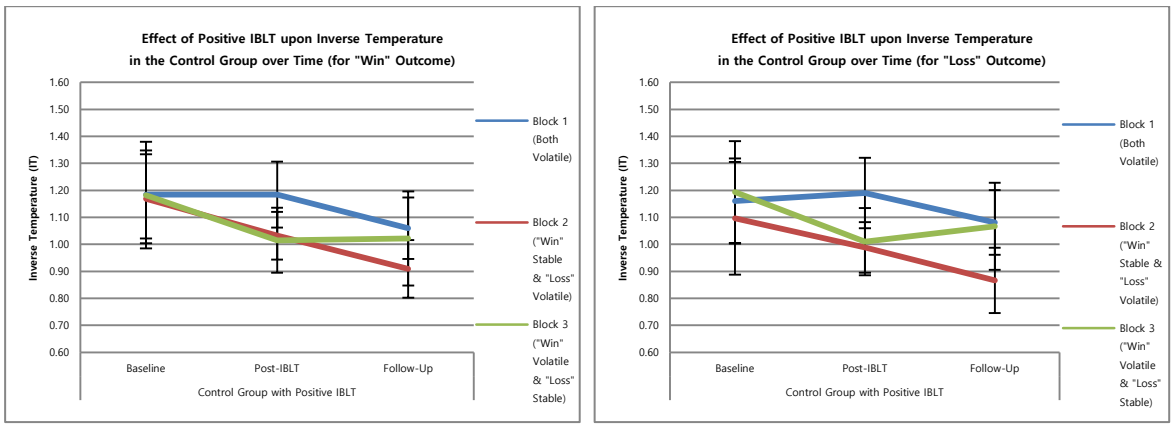


***Figure 13*. Effect of Positive IBLT upon Inverse Temperature in the Control Group for "Win" & "Loss" Outcomes over Time**

Figures display the mean (SEM) of inverse temperature across the study separately for each task block.

4. FERT

The various outcomes from the FERT task (accuracy, misclassification and reaction time) were been analysed separately. The FERT includes six emotions: anger, surprise, sad, happy, fear and disgusting. In order to simplify this, data was first analysed in terms of mean positive (happy & surprise) relative to mean negative (sad, fear, anger & disgust) emotion. This analysis provides a simple measure of affective bias. Following this exploratory analysis of the six individual emotions (*[4]) are reported for completeness.

***Table 3*. FERT Accuracy Changes**

	Previously Depressed Group with Neutral IBLT	Previously Depressed Group with Positive IBLT	Control Group with Neutral IBLT	Control Group with Positive IBLT
Negative Emotions				
Baseline	68.79 (3.98)	67.05 (5.35)	61.13 (19.42)	65.55 (9.79)
Post-IBLT	66.00 (5.38)	64.63 (9.36)	56.62 (19.18)	60.64 (11.20)
Follow-Up	67.79 (4.54)	64.13 (8.33)	64.22 (9.60)	61.64 (13.06)
Positive Emotions				
Baseline	60.17 (6.16)	59.39 (9.11)	59.57 (8.31)	61.86 (7.20)
Post-IBLT	63.76 (6.39)	58.77 (5.45)	58.86 (9.48)	61.76 (6.21)
Follow-Up	56.06 (10.96)	58.89 (10.18)	63.34 (7.08)	61.16 (5.70)
Anger				
Baseline	68.53 (8.28)	68.80 (10.54)	60.37 (18.35)	69.08 (12.77)
Post-IBLT	67.58 (7.54)	61.45 (13.12)	54.64 (22.63)	63.62 (12.65)
Follow-Up	69.37 (9.51)	63.83 (10.62)	67.27 (10.19)	66.67 (17.51)
Disgusting				
Baseline	69.85 (5.62)	68.05 (8.26)	64.42 (15.32)	59.35 (20.86)
Post-IBLT	63.73 (11.56)	68.60 (17.60)	58.49 (12.85)	54.97 (25.68)
Follow-Up	65.83 (11.90)	65.98 (21.06)	62.00 (12.96)	51.25 (27.33)
Fear				
Baseline	65.22 (9.28)	55.21 (11.92)	56.52 (22.26)	62.02 (10.17)
Post-IBLT	65.21 (8.15)	59.94 (12.08)	60.38 (21.92)	64.18 (11.43)
Follow-Up	69.86 (3.31)	59.65 (9.95)	63.65 (8.09)	68.64 (7.44)
Sad				
Baseline	71.60 (7.10)	75.89 (4.03)	63.40 (24.69)	71.58 (10.17)
Post-IBLT	67.50 (6.55)	68.46 (9.87)	53.20 (22.98)	59.77 (9.38)
Follow-Up	66.09 (9.41)	67.14 (6.84)	63.82 (16.24)	60.00 (7.91)
Happy				
Baseline	51.36 (8.53)	47.66 (12.63)	48.35 (14.33)	52.51 (10.90)
Post-IBLT	65.10 (8.60)	56.84 (10.06)	53.17 (17.67)	57.22 (8.24)
Follow-Up	55.80 (11.24)	58.88 (12.09)	56.29 (13.99)	57.08 (10.18)
Surprise				
Baseline	69.07 (9.02)	70.93 (8.33)	70.82 (6.19)	71.34 (6.78)
Post-IBLT	62.35 (9.45)	60.60 (4.53)	64.59 (3.66)	66.29 (6.77)
Follow-Up	56.32 (13.32)	58.98 (12.40)	70.36 (4.66)	65.24 (4.44)

Data are given as mean (standard deviation) unless indicated. ANOVA, Analysis of Variance; IBLT, Information Bias Learning Task.

[4] *Note*: The emotion 'neutral' is generally not analysed as a separate emotion.

***Table 4*. FERT Misclassification Changes**

	Previously Depressed Group with Neutral IBLT	Previously Depressed Group with Positive IBLT	Control Group with Neutral IBLT	Control Group with Positive IBLT
Negative Emotions				
Baseline	10.31 (9.99)	14.65 (11.04)	7.08 (7.02)	7.75 (4.31)
Post-IBLT	4.81 (3.83)	7.81 (5.80)	4.39 (5.19)	6.05 (7.45)
Follow-Up	13.05 (17.70)	7.20 (6.84)	4.29 (6.19)	5.02 (4.45)
Positive Emotions				
Baseline	1.22 (1.17)	2.81 (2.67)	3.28 (4.44)	3.70 (2.38)
Post-IBLT	2.44 (1.68)	2.61 (2.81)	2.92 (3.40)	2.74 (2.27)
Follow-Up	1.48 (0.62)	3.34 (2.61)	1.70 (1.56)	2.45 (1.87)
Anger				
Baseline	2.04 (1.02)	1.67 (1.63)	1.74 (1.84)	3.62 (3.31)
Post-IBLT	2.33 (1.59)	2.32 (2.72)	2.40 (2.50)	4.61 (4.97)
Follow-Up	2.71 (1.76)	2.04 (3.51)	3.98 (3.04)	5.16 (4.14)
Disgusting				
Baseline	4.38 (7.47)	3.10 (2.29)	2.03 (1.34)	1.26 (1.54)
Post-IBLT	2.06 (4.18)	1.38 (1.39)	1.36 (0.95)	1.30 (0.89)
Follow-Up	5.53 (11.80)	1.27 (1.37)	1.33 (1.15)	0.48 (0.37)
Fear				
Baseline	1.99 (1.89)	1.83 (1.94)	1.00 (1.34)	2.24 (3.14)
Post-IBLT	0.93 (0.82)	1.59 (0.69)	1.20 (0.98)	2.36 (3.84)
Follow-Up	1.27 (1.96)	1.92 (2.42)	0.38 (0.13)	2.30 (3.28)
Sad				
Baseline	2.83 (2.41)	6.93 (3.25)	3.51 (3.40)	4.70 (2.37)
Post-IBLT	2.43 (1.49)	6.72 (5.10)	2.43 (2.75)	4.03 (4.06)
Follow-Up	3.19 (2.46)	4.21 (3.74)	3.05 (3.96)	2.86 (2.50)
Happy				
Baseline	0.45 (0.41)	0.64 (0.92)	0.62 (0.78)	0.57 (0.32)
Post-IBLT	1.37 (1.17)	1.36 (1.62)	1.15 (1.08)	1.52 (1.23)
Follow-Up	0.95 (0.95)	1.79 (1.93)	0.72 (0.50)	1.11 (0.49)
Surprise				
Baseline	0.75 (0.63)	2.31 (2.46)	2.62 (3.29)	2.90 (2.24)
Post-IBLT	0.84 (0.50)	1.21 (1.88)	1.62 (2.51)	1.18 (1.32)
Follow-Up	0.56 (0.33)	1.47 (1.45)	0.93 (0.89)	1.27 (1.74)

Data are given as mean (standard deviation) unless indicated. ANOVA, Analysis of Variance; IBLT, Information Bias Learning Task.

***Table 5*. FERT Reaction Time Changes**

	Previously Depressed Group with Neutral IBLT	Previously Depressed Group with Positive IBLT	Control Group with Neutral IBLT	Control Group with Positive IBLT
Negative Emotions				
Baseline	1672.69 (196.30)	1905.25 (194.95)	1919.62 (321.58)	1801.01 (169.72)
Post-IBLT	1633.55 (224.13)	1884.81 (264.98)	1804.33 (175.84)	1687.39 (133.72)
Follow-Up	1630.94 (205.97)	1876.34 (330.56)	1666.46 (218.46)	1668.17 (195.55)
Positive Emotions				
Baseline	1587.07 (287.90)	1856.01 (205.88)	1631.47 (230.79)	1637.21 (149.09)
Post-IBLT	1507.16 (215.65)	1787.37 (192.45)	1526.58 (129.82)	1473.70 (101.57)
Follow-Up	1579.78 (195.74)	1819.01 (315.18)	1543.13 (182.21)	1499.69 (219.03)
Anger				
Baseline	1647.84 (227.29)	1898.47 (274.67)	1873.91 (454.36)	1708.50 (157.51)
Post-IBLT	1703.81 (329.78)	1973.69 (381.06)	1772.55 (216.67)	1667.37 (229.42)
Follow-Up	1624.89 (241.38)	2051.40 (585.80)	1736.18 (343.12)	1712.81 (335.23)
Disgusting				
Baseline	1693.50 (149.92)	1937.08 (344.69)	1929.52 (268.00)	2038.58 (239.28)
Post-IBLT	1709.36 (295.73)	1827.76 (328.22)	1832.79 (170.86)	1881.11 (493.63)
Follow-Up	1495.96 (126.92)	1807.83 (353.12)	1651.25 (197.44)	1756.25 (295.30)
Fear				
Baseline	1797.00 (299.43)	2239.11 (208.41)	2145.78 (274.27)	2044.94 (353.69)
Post-IBLT	1692.27 (200.78)	2185.28 (452.25)	2101.50 (474.66)	1874.21 (186.32)
Follow-Up	1868.45 (414.02)	2027.95 (312.73)	1877.44 (339.87)	1785.39 (237.47)
Sad				
Baseline	1575.61 (248.63)	1665.72 (184.26)	1858.73 (770.18)	1526.18 (143.84)
Post-IBLT	1451.27 (180.82)	1638.27 (210.07)	1512.65 (140.37)	1473.34 (142.42)
Follow-Up	1536.34 (251.05)	1685.69 (290.10)	1432.80 (213.69)	1445.43 (94.79)
Happy				
Baseline	1624.90 (278.83)	1794.31 (224.88)	1725.37 (267.43)	1643.59 (165.18)
Post-IBLT	1512.85 (225.39)	1785.66 (272.69)	1595.24 (169.26)	1490.12 (120.55)
Follow-Up	1523.02 (162.25)	1805.40 (306.06)	1541.20 (144.85)	1490.10 (173.42)
Surprise				
Baseline	1567.55 (311.73)	1893.31 (236.10)	1593.56 (256.86)	1626.27 (206.87)
Post-IBLT	1506.06 (213.42)	1798.26 (224.02)	1479.61 (121.72)	1464.52 (145.47)
Follow-Up	1640.54 (266.17)	1834.60 (341.09)	1560.32 (263.30)	1508.47 (267.18)

Data are given as mean (standard deviation) unless indicated. ANOVA, Analysis of Variance; IBLT, Information Bias Learning Task.

4.1. FERT Task: Baseline Comparison

- FERT Task: Baseline Accuracy

The interaction between emotion valence (negative vs. positive emotions) and group on accuracy was marginally significant with Greenhouse-Geisser correction, *Figure 14*; $F(1.00, 32.00)=3.65$, $p=0.065$. As can be seen from *Figure 14*, the previously depressed group displayed a negative bias relative to the control group, with a nonsignificantly higher accuracy for negative emotions (Previously Depressed M=67.92 & SD=4.64, Control M=63.34 & SD=15.09; $t(20.55)=1.22$, $p=0.24$) and a nonsignicantly lower accuracy for positive emotions (Previously Depressed M=59.78 & SD=7.52, Control M=60.71 & SD=7.63; $t(31.64)=-0.36$, $p=0.72$),. Paired-samples T-tests were conducted separately for each group, revealing significant difference between negative emotions (M=67.92, SD=4.64) and positive emotions in previously depressed group (M=59.78, SD=7.52); $t(15)=4.47$, $p<0.001$, with no difference between negative emotions (M=63.34, SD=15.09) and positive emotions (M=60.71, SD=7.63) in the control group; $t(17)=1.20$, $p=0.25$. In other words, the previously depressed group displayed an absolute negative bias, whereas the control group did not.

In further analysis, the omnibus interaction between the six separate emotions (anger, disgusting, fear, happy, sad & surprise) and group did not reach significance, $F(5, 160)=1.23$, $p=0.30$, although as can be seen from *Figure 15* the previously depressed group showed evidence for higher accuracy for sad, disgust and anger. Independent-samples T-tests were conducted to compare subject group difference for these emotions. However, no significant difference between groups was found ($p>0.14$). Overall, these results suggest the presence of a greater negative bias in face processing in the previously depressed relative to control group.

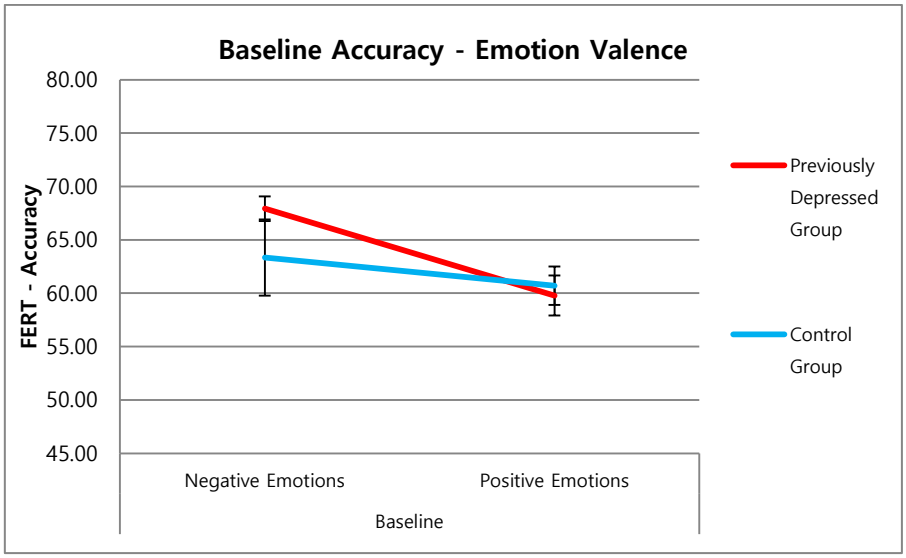


***Figure 14*. Baseline Accuracy in Negative & Positive Emotions**

Mean (SEM) of accuracy score of negative & positive emotions separately for the subject groups

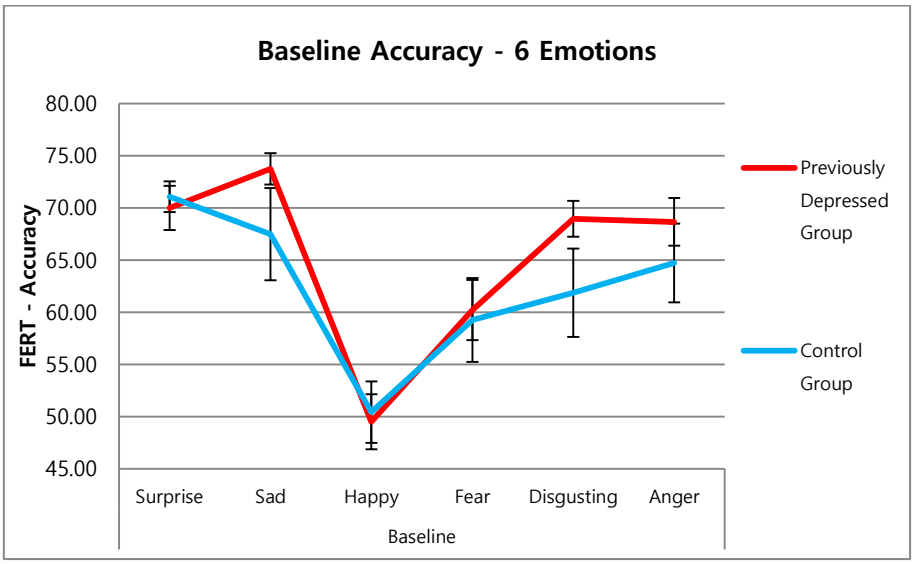


***Figure 15*. Baseline Accuracy in 6 Individual Emotions**

Mean (SEM) of accuracy score of 6 individual emotions separately for the subject groups

- FERT Task: Baseline Misclassification

The interaction of Emotional Valence (positive vs. negative) × Subject Group was significant, $F(1, 32)=4.92$, $p=0.03$. Again (*Figure 16*), this indicated a numerically greater negative bias in the previously depressed group, who misclassified more faces as negative and fewer faces as positive relative to the control group: First, independent-samples T-tests were conducted separately with negative emotions and positive emotions to assess the difference of FERT misclassification by subject group. There was a marginal difference between previously depressed group (M=12.48, SD=10.41) and control group (M=7.42, SD=5.66) for negative emotions; $t(22.54)=1.73$, $p=0.10$, with no difference between previously depressed group (M=2.01, SD=2.15) and control groups (M=3.49, SD=3.46) positive emotions; $t(28.81)=-1.51$, $p=0.14$. Second, paired-samples T-tests were conducted separately within each subject group to test for differences of FERT misclassification between negative and positive emotions. There was a significant difference between negative emotions (M=12.48, SD=10.41) and positive emotions (M=2.01, SD=2.15) in the previously depressed group; $t(15)=4.14$, $p=0.001$, as well as a significant difference between negative emotions (M=7.42, SD=5.66) and positive emotions (M=3.49, SD=3.46) in the control group; $t(17)=2.39$, $p=0.03$.

The interaction between the six individual emotions and subject group was again not significant, with Greenhouse-Geisser correction, $F(3.18, 101.75)=2.04$, $p=0.11$, but as can be seen from *Figure 17*, there was a tendency for the previously depressed group to show higher misclassification rates for sad and disgusted faces, although the subject group differences were not significant ($p>0.15$).

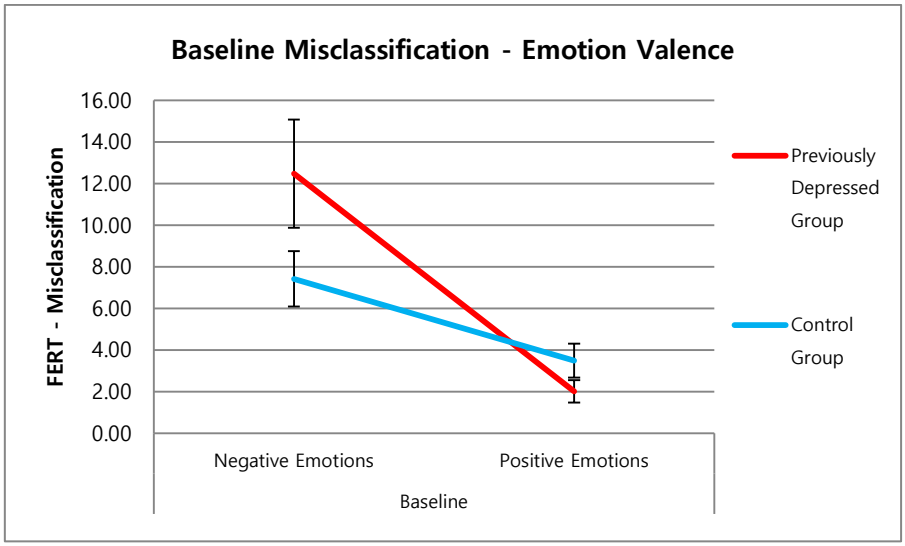


***Figure 16*. Baseline Misclassification in Negative & Positive Emotions**

Mean (SEM) of misclassification score of negative & positive emotions separately for the subject groups

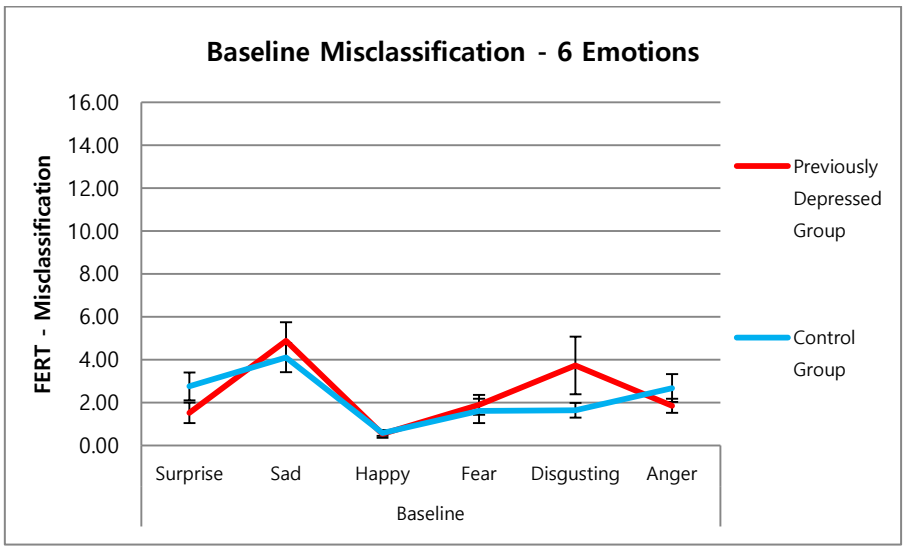


***Figure 17*. Baseline Misclassification in 6 Individual Emotions**

Mean (SEM) of misclassification score of 6 individual emotions separately for the subject groups

- FERT Task: Baseline Reaction Time

The interaction of Valence × Subject Group (*Figure 18*) was marginal with Greenhouse-Geisser correction, $F(1.00, 32.00)=3.04$, $p=0.09$, with the pattern of results again suggesting a greater negative bias for the previously depressed group than controls (i.e. faster for negative faces and slower for positive faces than the control group): An independent-samples T-test reported no significant difference between subject groups in both valence conditions ($p>0.30$), however, a paired-samples T-test reported a significant difference of reaction time between negative emotions (M=1860.31, SD=256.79) and positive emotions (M=1634.34, SD=188.50) in the control group; $t(17)=2.99$, $p=0.01$. The difference of reaction time between negative emotions (M=1788.97, SD=223.92) and positive emotions (M=1721.54, SD=278.84) in the previously depressed group was not significant; $t(15)=1.48$, $p=0.16$. The interaction of 6 Emotions × Subject Group was not significant with Greenhouse-Geisser correction, $F(3.50, 108.63)=1.96$, $p=0.12$.

As can be seen from *Figure 19*, previously depressed participants were marginally quicker to label faces as “disgusting” $t(32)=-1.83$, $p=0.08$, with other differences non-significant.

Overall, accuracy, misclassification and reaction time data from the FERT task are consistent with a greater negative bias in the previously depressed than control groups.

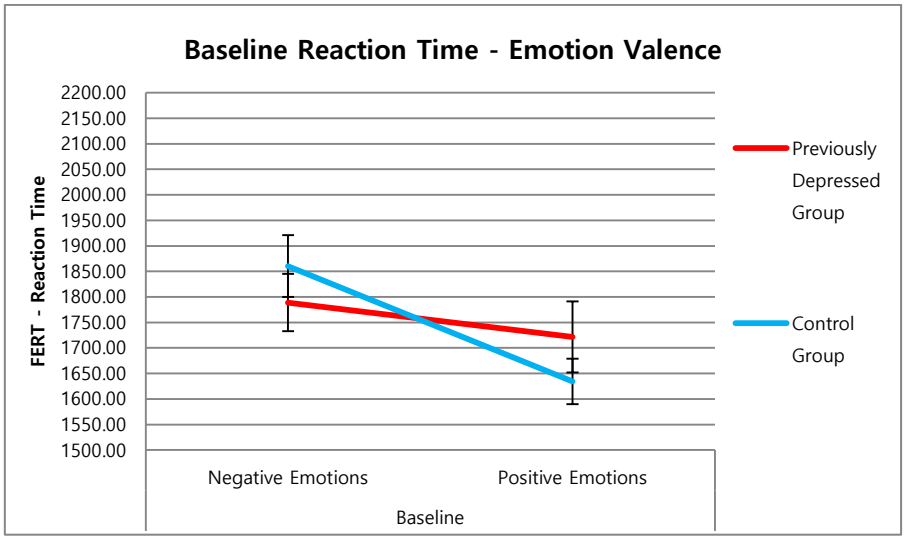


***Figure 18*. Baseline Reaction Time in Negative & Positive Emotions**

Mean (SEM) of reaction time of negative & positive emotions separately for the subject groups

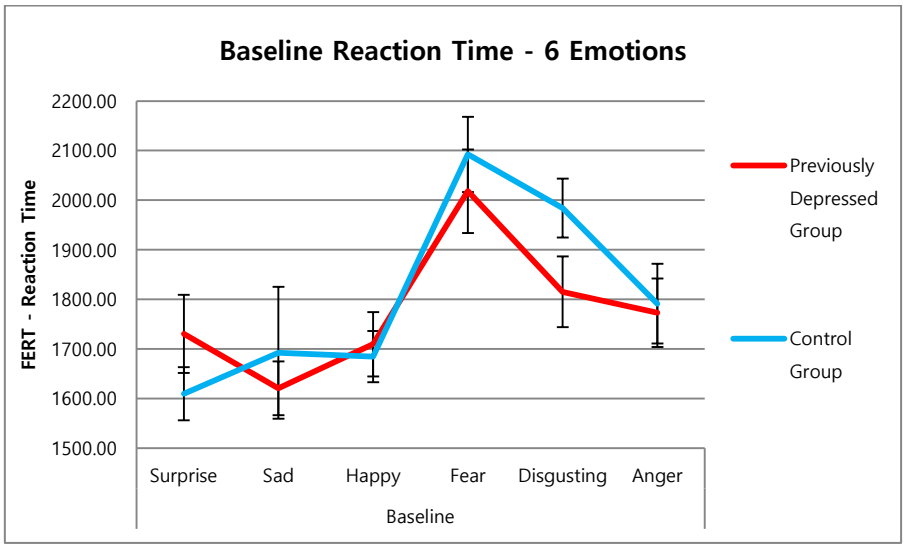


***Figure 19*. Baseline Reaction Time in 6 Individual Emotions**

Mean (SEM) of reaction time of 6 individual emotions separately for the subject groups

4.2. The Effect of IBLT on the FERT Task

- Accuracy

There was no significant effect of IBLT Type x Visit either across, $F(2, 44)=0.06$, $p=0.95$, or between emotional valence (IBLT Type x Visit x Valence), $F(2, 44)=0.29$, $p=0.75$. Furthermore, there was no interaction of these effects with subject group: p values for all interactions containing Subject Group was bigger than 0.13.

Overall, no convincing evidence for an effect of IBLT on FERT task accuracy was found.

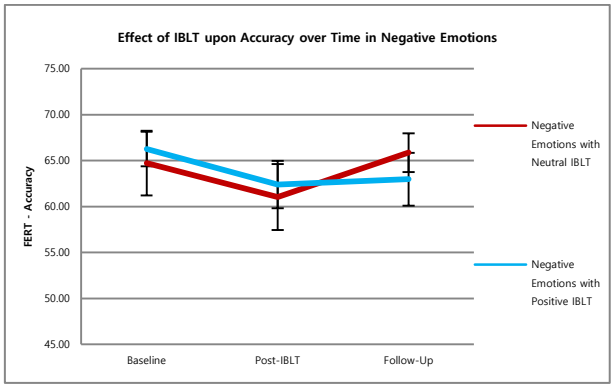


***Figure 20*. Effect of IBLT upon Accuracy over Time in Negative Emotions**

Figures display the mean (SEM) of accuracy score across the study for each testing visit.

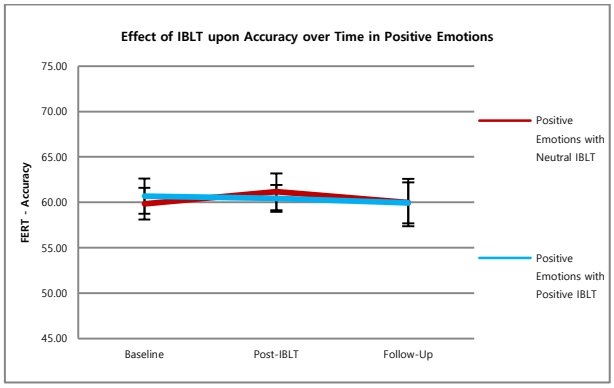


***Figure 21*. Effect of IBLT upon Accuracy over Time in Positive Emotions**

Figures display the mean (SEM) of accuracy score across the study for each testing visit.

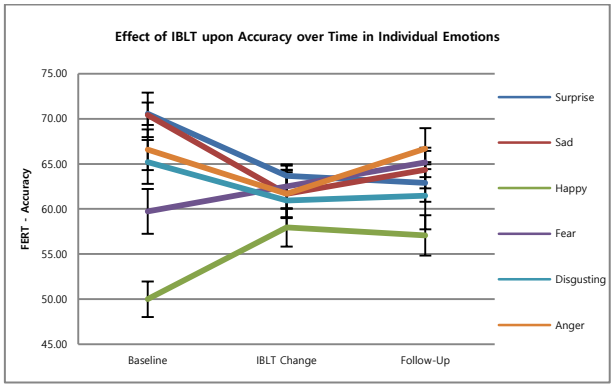


***Figure 22*. Effect of IBLT upon Accuracy over Time in Individual Emotions**

Figures display the mean (SEM) of accuracy score across the study separately for each of 6 individual emotions.

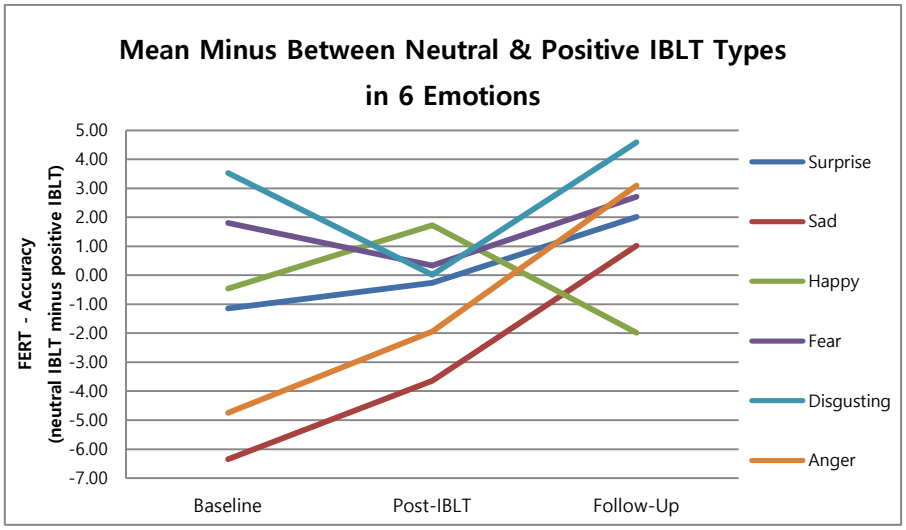


***Figure 23*. Relative Accuracy of the Neutral & Positive IBLT Groups for the 6 Emotions of the FERT Task**

Minus of mean (SEM) of accuracy scores in each of 6 individual emotions during 3 testing visits. In order to summarise data from all emotions and both IBLT groups this graph displays the difference between the mean FERT accuracy of participants in the neutral and positive IBLT groups, separately for each of the 6 emotions ("neutral IBLT – positive IBLT" of FERT Accuracy). If the value in the graph is negative, the neutral IBLT group is less accurate than the positive IBLT group. If the value is positive, then accuracy was higher in the neutral group.

- Misclassification

A repeated measures ANOVA reported a complex marginal effect of Visit × Valence × Subject Group × IBLT Type, $F(2, 44)=2.90$, $p=0.07$ (See *Figure 25* & *Figure 26*). However, analysing the 6 emotions separately, the effect of Visit × Emotion × Subject Group × IBLT Type was not significant with Greenhouse-Geisser correction, $F(5.66, 124.51)=0.79$, $p=0.57$.

Overall, the analysis of misclassifications in the FERT task tells that there is a significant effect of time (visit), but this was not found in further analysis with 6 emotions. This might come from the small sample size.

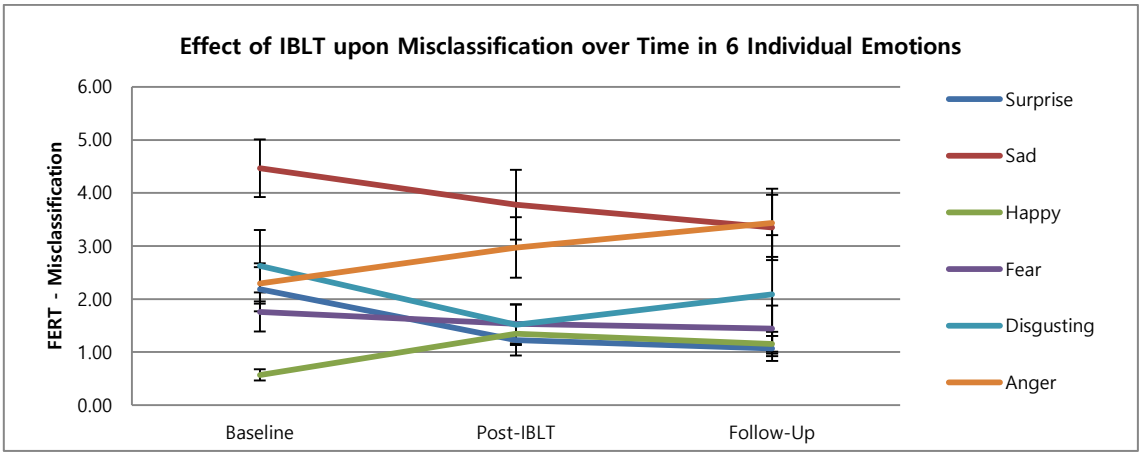


***Figure 24*. Effect of IBLT upon Misclassification over Time in 6 Individual Emotions**

Mean (SEM) of misclassification of 6 individual emotions across the study for each testing visit

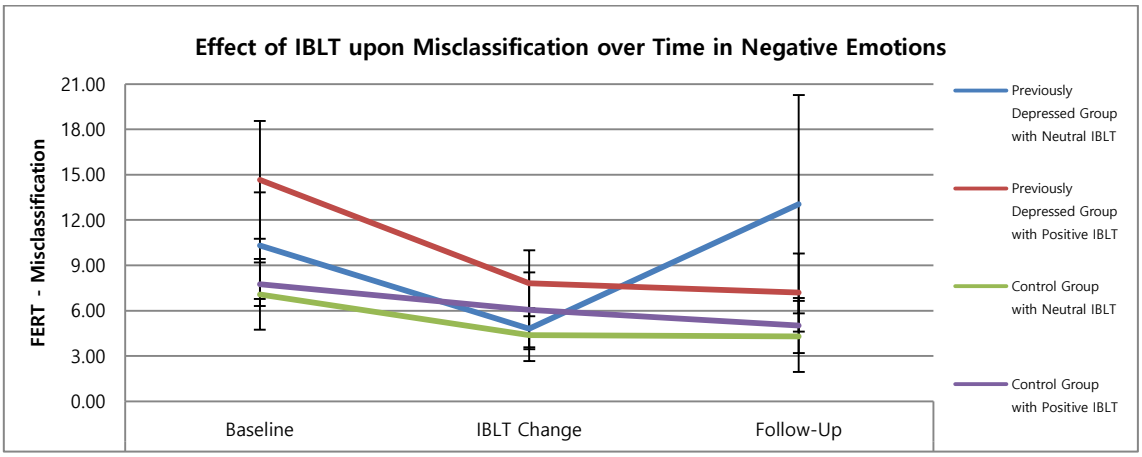


***Figure 25*. Effect of IBLT upon Misclassification over Time in Negative Emotions**

Mean (SEM) of misclassification of negative emotions across the study for each testing visit

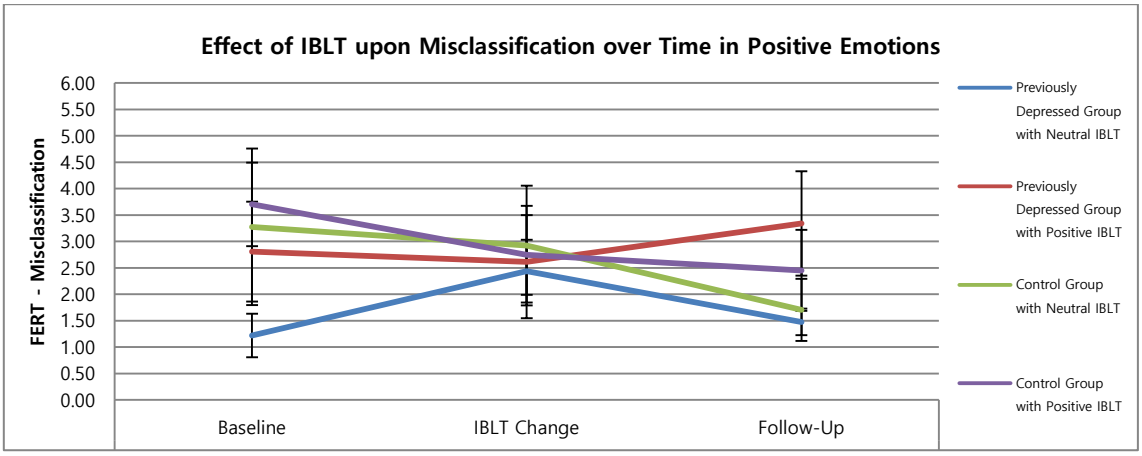


***Figure 26*. Effect of IBLT upon Misclassification over Time in Positive Emotions**

Mean (SEM) of misclassification of positive emotions across the study for each testing visit

- Reaction Time

The Visit × IBLT Type effect, $F(2, 44)=0.03$, $p=0.97$, the Visit × Subject Group × IBLT Type effect, $F(2, 44)=0.99$, $p=0.38$, the Visit × Valence × IBLT Type, $F(2, 44)=0.19$, $p=0.83$, and the Visit × Valence × Subject Group × IBLT Type effect, $F(2, 44)=0.21$, $p=0.81$, were nonsignificant. Analysing the emotions separately revealed no significant effect of IBLT on reaction time of Visit × Emotion × Subject Group × IBLT Type was not a significant with Greenhouse-Geisser correction, $F(5.92, 130.24)=1.31$, $p=0.26$, as well. All interactions containing Visit and IBLT Type were not significant ($p>0.26$).

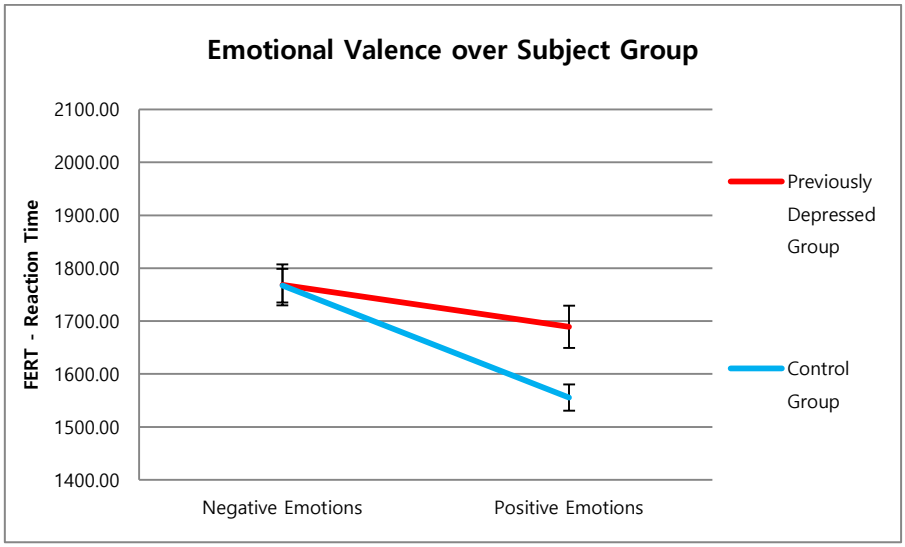


***Figure 27*. Interaction between Emotional Valence and Subject Group (Reaction Time)**

Mean (SEM) of reaction time of negative & positive emotions separately for each subject group

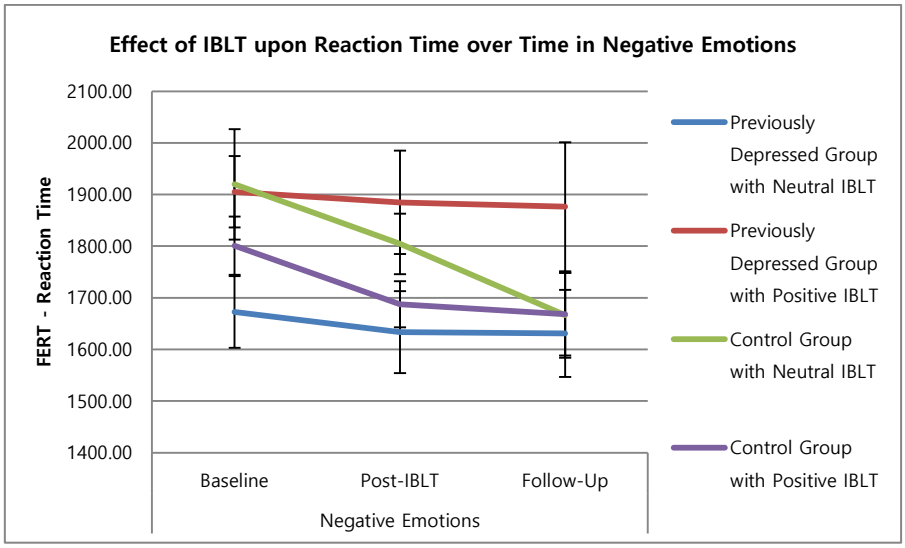


***Figure 28*. Effect of IBLT upon Reaction Time over Time in Negative Emotions**

Figure displays the mean (SEM) of reaction time in negative emotions across the study for each testing visit.

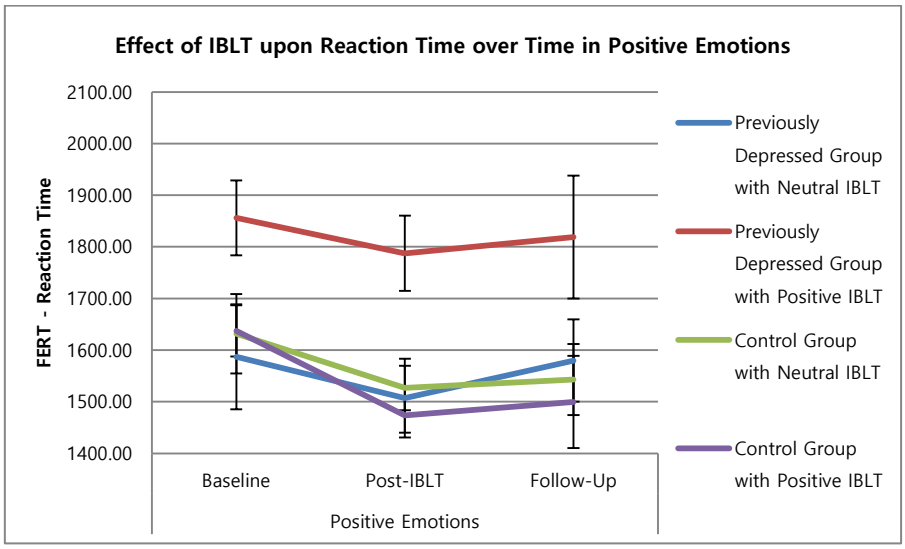


***Figure 29*. Effect of IBLT upon Reaction Time over Time in Positive Emotions**

Figure displays the mean (SEM) of reaction time in positive emotions across the study for each testing visit.

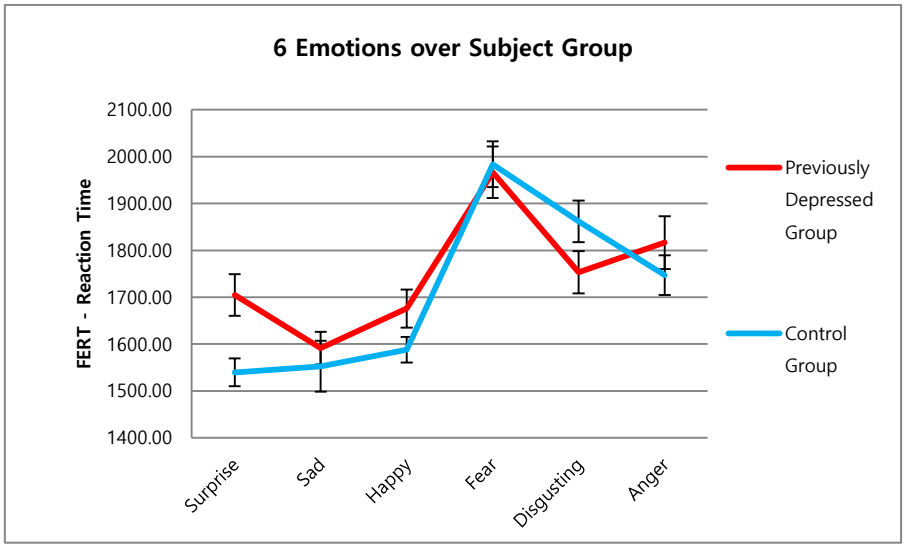


***Figure 30*. Interaction between 6 Individual Emotions & Subject Group (Reaction Time)**

Figure displays the mean (SEM) of reaction time of 6 individual emotions separately for each subject group.

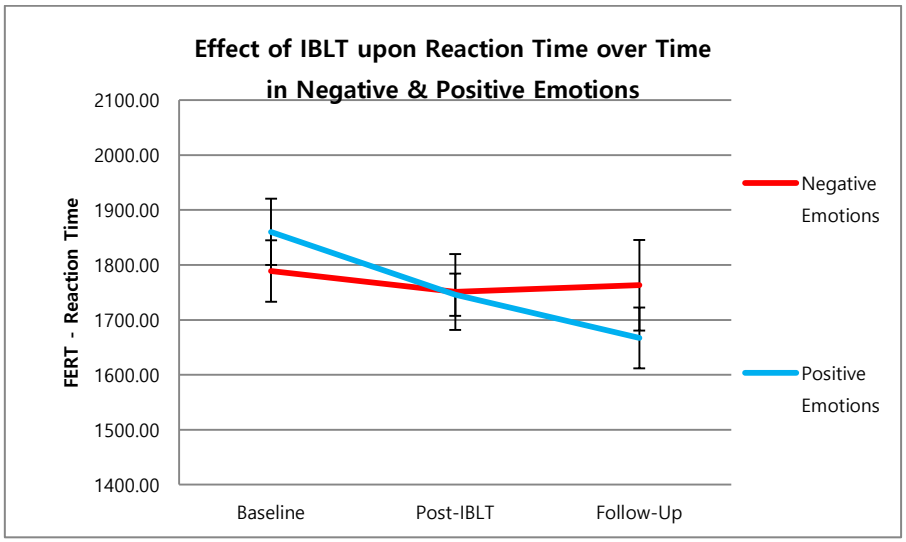


***Figure 31*. Effect of IBLT upon Reaction Time over Time in Negative & Positive Emotions**

Figure displays the mean (SEM) of reaction time in negative & positive emotions across the study for each testing visit.

5. Questionnaire Measures

Baseline differences in questionnaire scores are reported in *Table 2*. In this section, the effect of IBLT on these scores is reported. Where no significant effects are found the minimum p value for interactions which contain visit and IBLT group (i.e. with or without subject group) is reported for brevity. As groups differed on these measures at baseline and thus a larger effect of IBLT may be expected in the previously depressed group, follow up analyses of the effect of IBLT over time were run separately in the control and previously depressed groups.

6. QIDS

A repeated measures ANOVA reported no significant effect of any relevant interaction ("Visit", "Visit × IBLT Type", "Visit × Subject Group", "Visit × IBLT Type × Subject Group") (p>0.62) either across all groups or in each group separately.

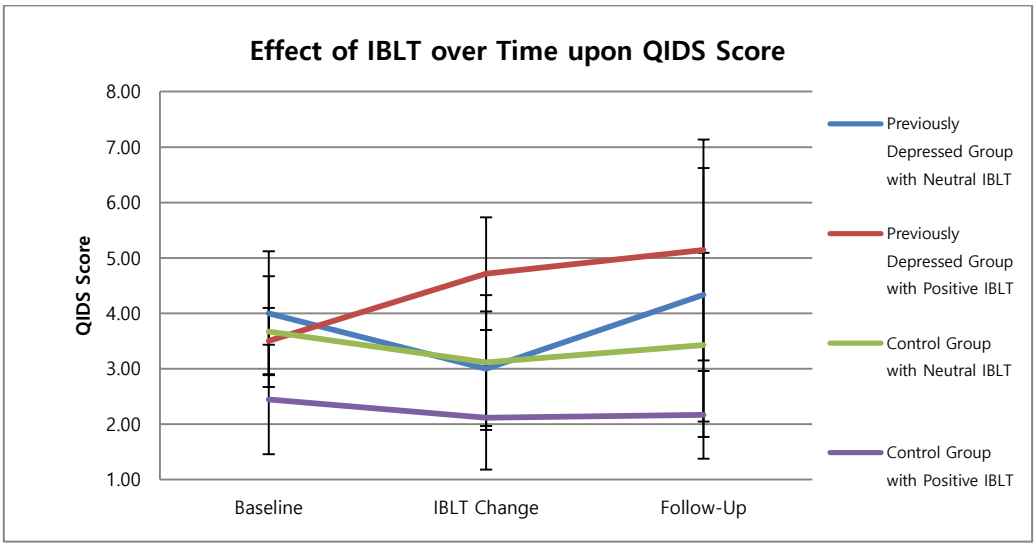


***Figure 32*. Effect of IBLT upon QIDS over Time**

Mean (SEM) of QIDS score of 4 groups across the study for each testing visit

7. SHAPS

A marginal effect of Visit × Subject Group × IBLT Type, with Greenhouse-Geisser correction, $F(1.43, 31.39)=3.26$, $p=0.07$ was found.

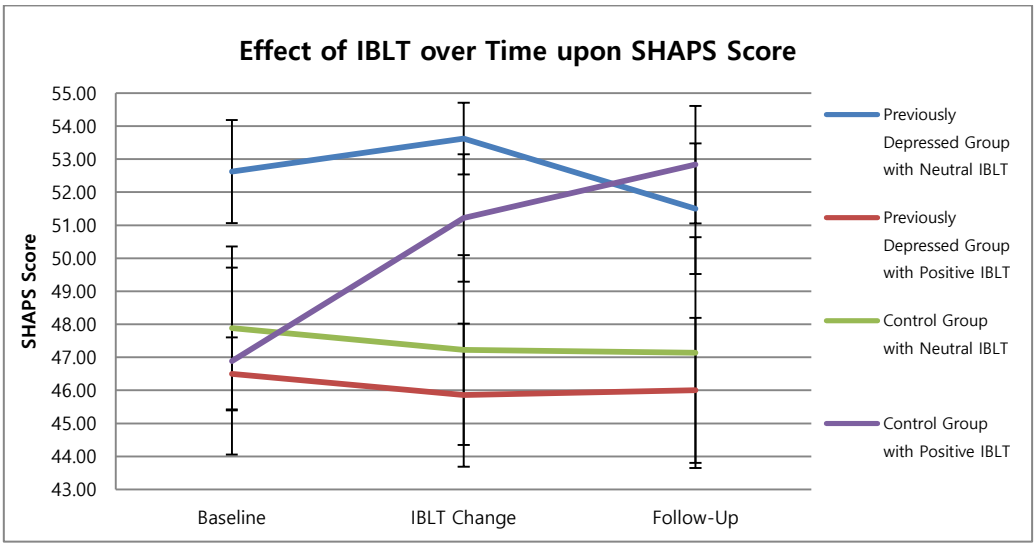


***Figure 33*. Effect of IBLT upon SHAPS over Time**

Mean (SEM) of SHAPS score of 4 groups across the study for each testing visit

The Visit × IBLT Type interaction, $F(2, 22)=0.60$, $p=0.56$, within the previously depressed group was not significant. However, the main effect of IBLT Type was significant, $F(1, 11)=6.31$, $p=0.03$, which, as illustrated in *Figure 33* above indicates a failure of randomisation, with a baseline difference in the positive and neutral IBLT groups at baseline being maintained throughout the study. However, the same analysis within controls reported a marginal effect of Visit × IBLT Type, with Greenhouse-Geisser correction, $F(1.19, 13.06)=3.45$, $p=0.08$, with the positively trained group showing an improvement in SHAPS score relative to the neutral group (i.e. showing less anhedonic symptoms). Independent-samples T-tests were conducted within controls to compare the

neutral and positive IBLT types across time points. However, none of these comparisons were significant (p>0.13). A paired-samples T-test was conducted to assess the change between testing visits of the two IBLT conditions within control subjects. Again, none of these comparisons were significant (p>0.12). Overall, there was some evidence of a beneficial effect of positive IBLT on symptoms of anhedonia only within the control group, although a number of the critical analyses were only significant at a trend level.

8. Trait-STAI

No effects of the explanatory variables or their interactions were apparent when analysing both groups together or separately suggesting that IBLT group did not differentially influence symptoms of trait anxiety over time.

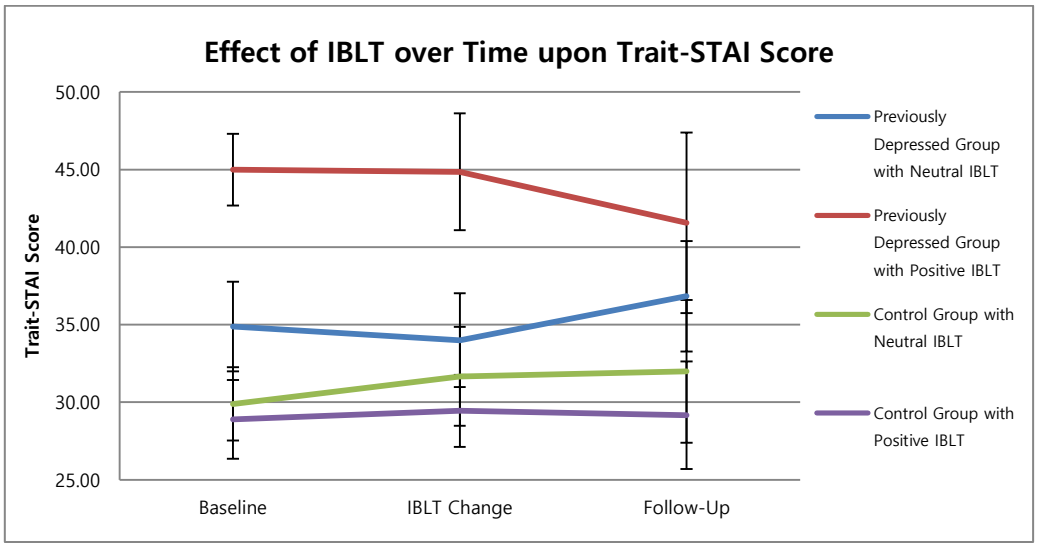


***Figure 34*. Effect of IBLT upon Trait-STAI over Time**

Mean (SEM) of Trait-STAI score of 4 groups across the study for each testing visit

9. DAQ

The analysis did not find any significant effect of learning training on scores of the DAQ either across or within groups.

Interestingly, the DAQ score of the previously depressed group who took positive IBLT is higher than that of the other three subject groups (see *Figure 35*), although the interaction between Subject Group × IBLT Type was not significant, $F(1, 22)=2.45$, $p=0.13$, this again looks like a failure of randomisation.

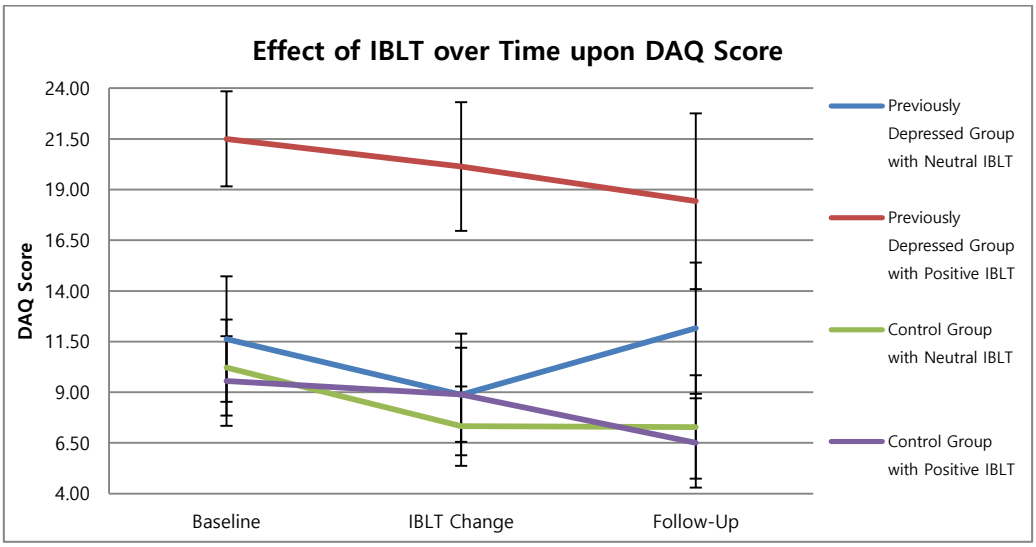


***Figure 35*. Effect of IBLT upon DAQ over Time**

Mean (SEM) of DAQ score of 4 groups across the study for each testing visit

10. RSS

No effect of IBLT upon participants' rumination was found either across or within groups.

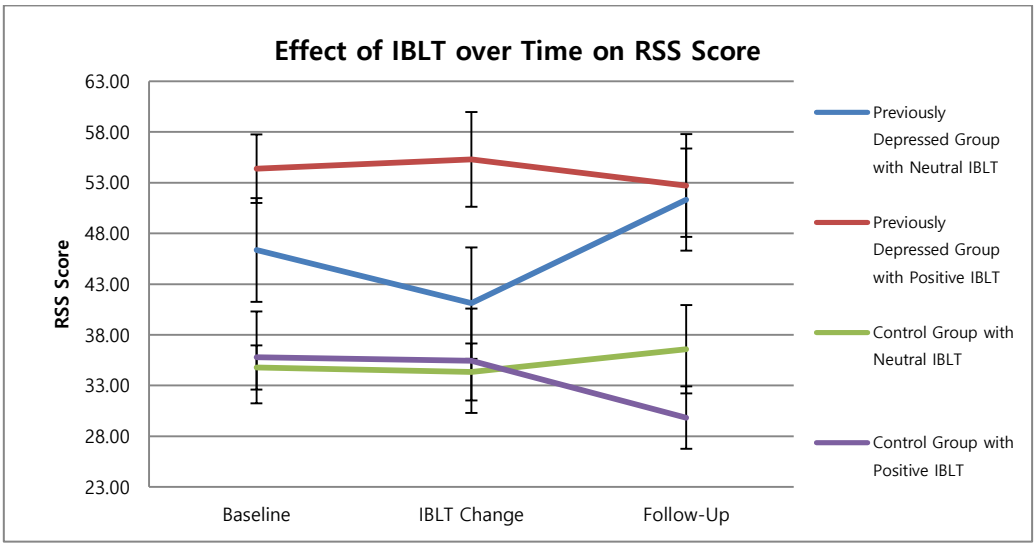


***Figure 36*. Effect of IBLT upon RSS over Time**

Mean (SEM) of RSS score of 4 groups across the study for each testing visit

Ⅳ. Discussion

The present study investigated baseline differences in learning parameters for positive and negative outcomes as well as other cognitive-affective measures. It also investigated the effect of information bias training on these cognitive measures and symptoms scores in previously depressed patients and control participants.

The main aims of the study were to (1) compare learning rates of previously depressed and control subjects for negative and positive events at baseline, (2) examine whether and how much the learning training (IBLT) modified these learning rates, and (3) assess the changes in the FERT and depression-related questionnaires induced by the IBLT.

Are There Baseline Differences Learning Rates, or Other Cognitive Measures, between Previously Depressed & Control Participants?

The study found little evidence to support the hypothesis that people with previous depression would show different learning rates for positive and negative events compared to control participants. The block design of the assessment task, which aimed to affect participants' learning rates by presenting them with variable volatilities of positive and negative outcomes, was successful in that, across participants, a higher learning rate was used for those outcomes which were volatile in a given block. This results is consistent with that reported previously using this task (Pulcu & Browning, 2017). However there was little evidence that

this effect differed in the previously depressed and control groups. Interestingly, a clear effect of previous depression was found in the FERT task with negative bias in the misclassification of emotional faces: Previously depressed individuals misclassified significantly more faces as negative than positive in the baseline assessment relative to controls (*Figure 16*). A similar pattern of results was also seen for the accuracy and reaction time data from the same task, although these effects did not all reach traditional levels of significance. Analyses of the separate groups indicated that the previously depressed individuals displayed a negative bias (accurately identifying more negative than positive faces) whereas never depressed individuals did not.

One might ask why a group difference between previously and never depressed individuals was found in the FERT task, but not in the IBAT task. First, the variability in these results may simply be a factor of the small sample size in the current study. The 34 participants included in this analysis, giving 16 and 18 per subject group, is clearly small and markedly reduces the power of the current study to detect effects should they be present. Thus it is possible that the negative baseline IBAT results arise due to a lack of power. Second, it is possible that the sensitivity of the tasks differs. The IBAT may simply be relatively less sensitive to group differences than the FERT task, making it hard to observe significant group differences. Third, the study focuses on comparison of previously depressed patients and controls. Excluding currently depressed subjects, may reduce the ability to detect state-like task effects. If future analysis of the full study data (which includes current depressed individuals) finds an effect on the IBAT task for the currently depressed rather than the previously depressed group, this would

indicate that the learning abnormalities measured by the task are state markers of depression (i.e. you only have them when you are actually depressed) rather than trait markers (you have them between episodes of depression). Last, it may simply be that depression (either current or previous) is simply not associated with differential learning in response to positive and negative outcomes, at least as measured by the IBAT task. The previous studies reporting learning effects in depression (Eshel & Roiser, 2010) provide some reassurance on this point, but the issue would best be addressed with a much larger study in which the association between depression and learning rate abnormalities can be clearly defined.

Does the IBLT Alter Learning Rate?

The second aim of the current study was to measure the impact of the IBLT training on cognitive measures—primarily on the learning rate parameter that the training was designed to alter. An effect of the IBLT task was seen for learning rates, although only at a trend level (Visit x Outcome x IBLT Type, p=0.06). This effect was in the expected direction with the change of learning rates over the 2 weeks of IBLT training differing for positive relative to neutral IBLT. Specifically, positive IBLT, in which positive outcomes were volatile and negative outcomes stable, led to an increase in learning rates only for positive outcomes, whereas neutral IBLT (in which both outcomes were volatile) led to an increase in learning rate for both outcomes. This demonstrates that the fundamental process targeted by IBLT is malleable, although analysis of data from the follow up test session suggests that the induced change in learning rate does erode after completion of

the training (*NB*: Similar results were reported by Browning, Holmes, Charles, Cowen & Harmer in 2012 using an attentional bias modification procedure.). Given the trend level effect described in this analysis, it will be important to ensure that this effect is also seen when the full study data set is available.

Does the IBLT Alter Other Learning Parameters?

The study also analysed inverse temperature, which estimates how randomly participants make decisions (or, alternatively, how much they are influenced by what they have learned). No baseline differences in inverse temperature were found between the control and previously depressed participants, but inverse temperature did change over the course of the study. The observed effect was relatively complex and, unlike the learning rate effect described above, difficult to neatly attribute to the effect of the IBLT training task. Post-hoc analysis of the inverse temperature data found that the effect of IBLT was more marked in the previously depressed group than control group. However, the pattern of effect even in previously depressed group was complex, with different effects on the different task blocks. More specifically, the inverse temperature applied increased for blocks in which the wins were volatile (Blocks 1 & 3 above) and decreased for blocks in which wins were stable (Block 2). Interpretation of this relatively complex interaction in the small sample of the current study must be cautious as it may clearly be a false positive. One possible interpretation of this finding is that the previously depressed patients may have found the blocks where wins were non informative frustrating, leading them to lose interest and perform more randomly. However, this effect was not predicted before the study

commenced, so it will be important to attempt to replicate it to test the possibility that it is a chance finding.

An interesting issue regarding IBLT is seen in the number of points won by participants during training. As noted in *Table 2* and *Figure 6*, subjects in the neutral IBLT groups won significantly more than subjects in the positive groups. This effect is expected as neutral training allow participants to learn from both the positive and negative stimuli while in the positive training participants can only learn from the positive stimuli, with the neutral stimuli being uninformative (meaning that, overall, participants will win less). One thing to consider is whether this imbalance may be avoided in future studies. One possibility would be to have a control condition in which losses were informative but wins were not (as in Block 2 of the IBAT task). As, in this case, participants would be able to learn from the negative but not the positive outcomes, one would expect performance to be roughly the same as that for the positive group in the current study. This approach was not employed in the current study because a negative training condition would be predicted to make people more depressed which is not ethical, particularly in clinical groups such as recruited here. An alternative approach to reduce the difference in performance between conditions would be to use a probability other than 50% for occurrences of the outcomes in the stable condition. For example, if the outcome in the stable condition was associated with one of the stimuli with a probability of 75%, it would be possible to learn this contingency and use it when making a decision. While it would be important in such a task to ensure that the volatile outcome was still able to drive behaviour such that subjects sometimes chose the option not favoured by the stable outcome, the prediction is that

participants would still favour learning from the volatile outcome and so it may be possible to induce a learning bias in the absence of such a large difference in performance between training conditions.

Does the Effect of the IBLT Generalise to Other Cognitive & Symptoms Measures?

Having established a relatively clear learning effect of the basic training intervention on participants' learning rates, the next question is the degree to which this effect generalises to the other outcome measures. In the current study we tested this using a separate cognitive measure (the FERT task) as well as a range of questionnaire measures of symptoms and cognitive processes. Here there was little evidence of generalisation, with no clear separation between IBLT groups following training. The one possible exception to this was for the SHAPS questionnaire in which a significant effect of IBLT was found within the control group only. This effect was in the predicted direction, with SHAPS scores (which are higher as one experiences more pleasure in daily activities) increasing more in the positively than neutral training group. The confidence in this result is somewhat reduced as a similar pattern was not found in the previously depressed group, although here there was a large baseline difference between IBLT groups likely resulting from a failure of randomisation within this group. Given the interim nature of this analysis it is difficult to draw strong conclusions from these generally negative findings (or from the occasional possible positive finding), in particular it is important to note that the sample size included in this analysis is small, and so the current study is not well powered to detect any effect of training, and the population was limited to participants who are not currently depressed (*NB*:

although currently depressed patients are being recruited to the study, too few had completed it to be included in this analysis) which will reduce the sensitivity of detecting changes in symptom based scores. These considerations aside, it is important to acknowledge the possibility that the cognitive process targeted here (the estimated information content of positive and negative outcomes) may simply not be causally related to symptoms of depression and thus modifying it may not lead to any appreciable change in these symptoms. The current study, which is continuing to recruit patients should provide a more definitive assessment of these possibilities when it has been completed.

Study Limitations

There are important limitations to the current study: As discussed above, the sample size of the study was modest (8 or 9 participants per group), and the study analysed only previously depressed and control participants, but not currently depressed patients. This factor limits the confidence in positive results and the interpretability of the negative results. To some extent this limitation will be addressed by the ongoing recruitment for the study. However, even the final target samples size of the study is relatively modest (20 per group) and thus it would be useful to consider a larger study, particularly when looking at case-control baseline differences. Recently a number of influential studies (e.g. Gillan, Kosinski, Whelan, Phelps & Daw, 2016) have dealt with this problem using online platforms to deliver cognitive tasks to large numbers of participants. While the ability to clinically characterise participants in such studies is compromised, this may be more than

made up for by the very large sample sizes which may be recruited. A second limitation relates to the exclusion criteria of the study. Patients with depression which was associated with significant suicidality and those receiving treatment(s) were excluded from the study which reduces how representative the current sample is of all people who are depressed. These exclusions were incorporated in the study design to ensure a more homogenous population in this initial study (e.g. there is clear evidence that antidepressant treatment influences performance of the FERT task (Harmer et al., 2009, p. 1180-1183)), however one effect of these exclusions is that the population of patients is likely to have a relatively mild illness.

Future Work

The results described above suggest a number of ways in which the work reported in this thesis may be extended in the future. In addition to recruiting more participants and including currently depressed patients in the analysis, it would be interesting to test which task parameters optimised the effect of the IBLT training. For example, in the current study training occurred every day for 2 weeks, with participants completing one training session, which lasted 10 minutes, per day. These parameters were selected based on the estimated tolerance of depressed patients for relatively repetitive cognitive testing, however it may be that more frequent or longer (in terms of individual training sessions or duration of the whole training program) training may be more effective at inducing a generalised bias (Hallion & Ruscio, 2011). Second, different structures for the positive IBLT training may be useful. For example, as flexibility is important in learning, it may be useful

to have training which is predominantly rather than completely positive, with a structure such as “Block 3 (volatile positive event & stable negative event) → Block 2 (the opposite events compared to those in Block 3) → Block 3”. Such a regime would ensure participants had to adapt to constantly changing volatility while still maintaining a bias for positive information. Additionally, the current study has focussed on depression as this illness has regularly been associated with negative biases (Mathews & Macleod, 2005; Eshel & Roiser, 2010). However, previous work using different forms of bias modification (reviewed by Hallion & Ruscio, 2011) suggested that patients with anxiety were more responsive to such interventions. A difficulty in adapting learning to the volatility of outcomes has previously been described in anxious individual (Browning et al. 2015) and so it would be interesting to test the impact of a learning bias training, such as that used here, in an anxious group.

Conclusion

In conclusion, the current study provided evidence that a novel form of cognitive training, the IBLT, was able to modify the learning rates of previously depressed and control participants although did not find convincing evidence that doing so is helpful for depression. Overall the approach used in this study—identifying novel cognitive processes which may be important in the aetiology of depression and then assessing the impact of modifying those processes—promises to inform the development of interesting new approaches to treatment, although it has yet to be proven whether the specific methods tested in this study are useful.

REFERENCE

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

Andrade, L., Caraveo-anduaga, J. J., Berglund, P., Bijl, R. V., Graaf, R. D., Vollebergh, W., ... & Kawakami, N. (2003). The epidemiology of major depressive episodes: results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys. *International Journal of Methods in Psychiatric Research*, *12*(1), 3-21. DOI: 10.1002/mpr.138

Angermeyer, M. C., Holzinger, A., Matschinger, H., & Stengler-Wenzke, K. (2002). Depression and Quality of Life: Results of a Follow-Up Study. *International Journal of Social Psychiatry*, *48*(3), 189-199. DOI: 10.1177/002076402128783235

Beats, B. C., Sahakian, B. J., & Levy, R. (1996). Cognitive performance in tests sensitive to frontal lobe dysfunction in the elderly depressed. *Psychological Medicine*, *26*(3), 591-603. DOI: 10.1017/S0033291700035662

Beck, A. T. (1967). *Depression: Clinical, Experimental, and Theoretical Aspects*. New York, NY: University of Pennsylvania Press.

Beck, A. T. (1976). *Cognitive Therapy and the Emotional Disorders*. New York, NY: International Universities Press, Inc.

Beck, A. T., Rush, A. J., Shaw, B. F. & Emery, G. (1979). *Cognitive Therapy of Depression*. New York, NY: Guilford Press.

Behrens, T. E., Hunt, L. T., Woolrich, M. W. & Rushworth, M. F. (2008). Associative learning of social value. *Nature*, *456*(7219), 245-249. DOI: 10.1038/nature07538

Behrens, T. E., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214-1221. DOI: 10.1038/nn1954

Blackwell, S. E., Browning, M., Mathews, A., Pictet, A., Welch, J., Davies, J., ... & Holmes, E. A. (2015). Positive Imagery-Based Cognitive Bias Modification as a Web-Based Treatment Tool for Depressed Adults: A Randomized Controlled Trial. *Clinical Psychological Science*, *3*(1), 91-111. DOI: 10.1177/2167702614560746

Browning, M., Behrens, T. E., Jocham, G., O'reilly, J. X., & Bishop, S. J. (2015). Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nature Neuroscience*, *18*(4), 590-596. DOI: 10.1038/nn.3961

Browning, M., Blackwell, S. E., & Holmes, E. A. (2012). The Use of Cognitive Bias Modification and Imagery in the Understanding and Treatment of Depression. In *Behavioral Neurobiology of Depression and Its Treatment* (pp. 243-260). Springer Berlin Heidelberg. DOI: 10.1007/7854_2012_212

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. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(1), 8-20. DOI: 10.3758/CABN.10.1.8

Browning, M., Holmes, E. A., Charles, M., Cowen, P. J., & Harmer, C. J. (2012).

Using Attentional Bias Modification as a Cognitive Vaccine Against Depression. *Biological Psychiatry*, *72*(7), 572-579. DOI: 10.1016/j.biopsych.2012.04.014

Casacalenda, N., Perry, J. C., & Looper, K. (2002). Remission in Major Depressive Disorder: A Comparison of Pharmacotherapy, Psychotherapy, and Control Conditions. *American Journal of Psychiatry*, *159*(8), 1354-1360. DOI: 10.1176/appi.ajp.159.8.1354

Christopher, G., & MacDonald, J. (2005). The impact of clinical depression on working memory. *Cognitive Neuropsychiatry*, *10*(5), 379-399. DOI: 10.1080/13546800444000128

Cox, B. J., Enns, M. W., & Taylor, S. (2001). The effect of rumination as a mediator of elevated anxiety sensitivity in major depression. *Cognitive Therapy & Research*, *25*(5), 525-534. DOI: 10.1023/A:1005580518671

Cristea, I. A., Kok, R. N., & Cuijpers, P. (2015). Efficacy of cognitive bias modification interventions in anxiety and depression: meta-analysis. *The British Journal of Psychiatry*, *206*(1), 7-16. DOI: 10.1192/bjp.bp.114.146761

Danion, J. M., Willard-Schroeder, D., Zimmermann, M. A., Grange, D., Schlienger, J. L., & Singer, L. (1991). Explicit Memory and Repetition Priming in Depression: Preliminary Findings. *Archives of General Psychiatry*, *48*(8), 707-711. DOI: 10.1001/archpsyc.1991.01810320031005

Dichter, G. S., Smoski, M. J., Kampov-Polevoy, A. B., Gallop, R., & Garbutt, J. C. (2010). Unipolar Depression Does Not Moderate Responses to the Sweet Taste Test. *Depression and Anxiety*, *27*(9), 859-863. DOI: 10.1002/da.20690

Dotson, V. M., Resnick, S. M., & Zonderman, A. B. (2008). Differential Association of Concurrent, Baseline, and Average Depressive Symptoms With Cognitive Decline in Older Adults. *The American Journal of Geriatric Psychiatry*, *16*(4), 318-330. DOI: 10.1097/JGP.0b013e3181662a9c

Driessen, E., & Hollon, S. D. (2010). Cognitive Behavioral Therapy for Mood Disorders: Efficacy, Moderators and Mediators. *Psychiatric Clinics of North America*, *33*(3), 537-555. DOI: 10.1016/j.psc.2010.04.005

Eshel, N., & Roiser, J. P. (2010). Reward and Punishment Processing in Depression. *Biological Psychiatry*, *68*(2), 118-124. DOI: 10.1016/j.biopsych.2010.01.027

Ekman, P., & Friesen, W. V. (1971). Constants Across Cultures in the Face and Emotion. *Journal of Personality and Social Psychology*, *17*(2), 124. DOI: 10.1037/h0030377

Elliott, C. L., & Greene, R. L. (1992). Clinical Depression and Implicit Memory. *Journal of Abnormal Psychology*, *101*(3), 572.

Elliott, R., Sahakian, B. J., Herrod, J. J., Robbins, T. W., & Paykel, E. S. (1997). Abnormal response to negative feedback in unipolar depression: evidence for a diagnosis specific impairment. *Journal of Neurology, Neurosurgery & Psychiatry*, *63*(1), 74-82. DOI: 10.1136/jnnp.63.1.74

Frank, E., Kupfer, D. J., Perel, J. M., Cornes, C., Jarrett, D. B., Mallinger, A. G., ... & Grochocinski, V. J. (1990). Three-year outcomes for maintenance therapies in recurrent depression. *Archives of General Psychiatry*, *47*(12), 1093-1099.

DOI: 10.1001/archpsyc.1990.01810240013002

Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *Elife*, *5*, e11305. DOI: 10.7554/eLife.11305

Gonzales, L. R., Lewinsohn, P. M., & Clarke, G. N. (1985). Longitudinal Follow-Up of Unipolar Depressives: An Investigation of Predictors of Relapse. *Journal of Consulting and Clinical Psychology*, *53*(4), 461-469. DOI: 10.1037/0022-006X.53.4.461

Gorwood, P., Corruble, E., Falissard, B., & Goodwin, G. M. (2008). Toxic Effects of Depression on Brain Function: Impairment of Delayed Recall and the Cumulative Length of Depressive Disorder in a Large Sample of Depressed Outpatients. *American Journal of Psychiatry*, *165*(6), 731-739. DOI: 10.1176/appi.ajp.2008.07040574

Gotlib, I. H., Krasnoperova, E., Yue, D. N., & Joormann, J. (2004). Attentional Biases for Negative Interpersonal Stimuli in Clinical Depression. *Journal of Abnormal Psychology*, *113*(1), 127. DOI: 10.1037/0021-843X.113.1.121

Hallion, L. S., & Ruscio, A. M. (2011). A Meta-Analysis of the Effect of Cognitive Bias Modification on Anxiety and Depression. *Psychological Bulletin*, *137*(6), 940-958. DOI: 10.1037/a0024355

Harmer, C. J., Shelley, N. C., Cowen, P. J., & Goodwin, G. M. (2004). Increased Positive Versus Negative Affective Perception and Memory in Healthy Volunteers Following Selective Serotonin and Norepinephrine Reuptake

Inhibition. *American Journal of Psychiatry*, *161*(7), 1256-1263. DOI: 10.1176/appi.ajp.161.7.1256

Harmer, C. J., Bhagwagar, Z., Perrett, D. I., Völlm, B. A., Cowen, P. J., & Goodwin, G. M. (2003a). Acute SSRI Administration Affects the Processing of Social Cues in Healthy Volunteers. *Neuropsychopharmacology*, *28*(1), 148-152. DOI: 10.1038/sj.npp.1300004

Harmer, C. J., Hill, S. A., Taylor, M. J., Cowen, P. J., & Goodwin, G. M. (2003b). Toward a Neuropsychological Theory of Antidepressant Drug Action: Increase in Positive Emotional Bias After Potentiation of Norepinephrine Activity. *American Journal of Psychiatry*, *160*(5), 990-992. DOI: 10.1176/appi.ajp.160.5.990

Harmer, C. J., O'Sullivan, U., Favaron, E., Massey-Chase, R., Ayres, R., Reinecke, A., ... & Cowen, P. J. (2009). Effect of Acute Antidepressant Administration on Negative Affective Bias in Depressed Patients. *American Journal of Psychiatry*, *166*(10), 1178-1184. DOI: 10.1176/appi.ajp.2009.09020149

Hecht, H., von Zerssen, D., & Wittchen, H. U. (1990). Anxiety and depression in a community sample: the influence of comorbidity on social functioning. *Journal of Affective Disorders*, *18*(2), 137-144. DOI: 10.1016/0165-0327(90)90070-O

Huys, Q. J., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biology of Mood & Anxiety Disorders*, *3*(1), 12. DOI: 10.1186/2045-5380-3-12

Jenkins, W., & McDowall, J. (2001). Implicit memory and depression: An analysis

of perceptual and conceptual processes. *Cognition and Emotion*, *15*(6), 803-812. DOI: 10.1080/02699930143000220

Julian, L. J. (2011). Measures of Anxiety: State-Trait Anxiety Inventory (STAI), Beck Anxiety Inventory (BAI), and Hospital Anxiety and Depression Scale-Anxiety (HADS-A). *Arthritis Care and Research*, *63*(S11), S467-S472. DOI: 10.1002/acr.20561

Keller, M. B., Lavori, P. W., Lewis, C. E., & Klerman, G. L. (1983). Predictors of Relapse in Major Depressive Disorder. *JAMA*, *250*(24), 3299-3304. DOI: 10.1001/jama.1983.03340240025024

Keller, M. B., Shapiro, R. W., Lavori, P. W., & Wolfe, N. (1982). Relapse in Major Depressive Disorder: Analysis With the Life Table. *Archives of General Psychiatry*, *39*(8), 911-915. DOI: 10.1001/archpsyc.1982.04290080031005

Kleim, B., Gonzalo, D. & Ehlers, A. (2011). The Depressive Attributions Questionnaire (DAQ): Development of a Short Self-Report Measure of Depressogenic Attributions. *Journal of Psychopathology & Behavioral Assessment*, *33*, 375. DOI: 10.1007/s10862-011-9234-9

Liu, H., Li, X., Han, B., & Liu, X. (2017). Effects of cognitive bias modification on social anxiety: A meta-analysis. *PloS One*, *12*(4), e0175107. DOI: 10.1371/journal.pone.0175107

Maj, M., Veltro, F., Pirozzi, R., Lobrace, S., & Magliano, L. (1992). Pattern of Recurrence of Illness After Recovery From an Episode of Major Depression: A Prospective Study. *American Journal of Psychiatry*, *149*(6), 795-800. DOI:

10.1176/ajp.149.6.795

Mathews, A., & MacLeod, C. (2005). Cognitive Vulnerability to Emotional Disorders. *Annual Review of Clinical Psychology*, *1*, 167-195. DOI: 10.1146/annurev.clinpsy.1.102803.143916

Matt, G. E., Vázquez, C., & Campbell, W. K. (1992). Mood-Congruent Recall of Affectively Toned Stimuli: A Meta-Analytic Review. *Clinical Psychology Review*, *12*(2), 227-255. DOI: 10.1016/0272-7358(92)90116-P

Mayberg, H. S. (2002). Mapping mood: An evolving emphasis on frontal-limbic interactions. In: Stuss, D.T., Knight, R.T. (Eds.), *Principles of Frontal Lobe Function*. Oxford University Press, Oxford, pp. 376-391.

McDermott, L. M., & Ebmeier, K. P. (2009). A meta-analysis of depression severity and cognitive function. *Journal of Affective Disorders*, *119*(1), 1-8. DOI: 10.1016/j.jad.2009.04.022

McFarland, B. R., & Klein, D. N. (2009). Emotional reactivity in depression: diminished responsiveness to anticipated reward but not to anticipated punishment or to nonreward or avoidance. *Depression and Anxiety*, *26*(2), 117-122. DOI: 10.1002/da.20513

Montgomery, S. A., & Asberg, M. A. R. I. E. (1979). A New Depression Scale Designed to be Sensitive to Change. *The British Journal of Psychiatry*, *134*(4), 382-389. DOI: 10.1192/bjp.134.4.382

Moussavi, S., Chatterji, S., Verdes, E., Tandon, A., Patel, V., & Ustun, B. (2007). Depression, chronic diseases, and decrements in health: results from the

World Health Surveys. *The Lancet*, *370*(9590), 851-858. DOI: 10.1016/S0140-6736(07)61415-9

Must, A., Szabó, Z., Bódi, N., Szász, A., Janka, Z., & Kéri, S. (2006). Sensitivity to reward and punishment and the prefrontal cortex in major depression. *Journal of Affective Disorders*, *90*(2), 209-215. DOI: 10.1016/j.jad.2005.12.005

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience*, *15*(7), 1040-1046. DOI: 10.1038/nn.3130

Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology*, *109*(3), 504-511. DOI: 10.1037/0021-843X.109.3.504

Nolen-Hoeksema, S., & Morrow, J. (1991). A prospective study of depression and post-traumatic stress symptoms following a natural disaster: The 1989 Loma Prieta Earthquake. *Journal of Personality & Social Psychology*, *61*(3), 115-121. DOI: 10.1037/0022-3514.61.1.115

Nolen-Hoeksema, S., Parker, L. E., & Larson, J. (1994). Ruminative Coping with Depressed Mood Following Loss. *Journal of Personality and Social Psychology*, *67*(1), 92. DOI: 10.1037/0022-3514.67.1.92

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking Rumination. *Perspectives on Psychological Science*, *3*(5), 400-424. DOI: 10.1111/j.1745-6924.2008.00088.x

O'neil, A., Sanderson, K., Oldenburg, B., & Taylor, C. B. (2011). Impact of

Depression Treatment on Mental and Physical Health-Related Quality of Life of Cardiac Patients: A Meta-Analysis. *Journal of Cardiopulmonary Rehabilitation and Prevention*, *31*(3), 146-156. DOI: 10.1097/HCR.0b013e3181fc0985

Pizzagalli, D. A., Iosifescu, D., Hallett, L. A., Ratner, K. G., & Fava, M. (2008). Reduced hedonic capacity in major depressive disorder: Evidence from a probabilistic reward task. *Journal of Psychiatric Research*, *43*(1), 76-87. DOI: 10.1016/j.jpsychires.2008.03.001

Pizzagalli, D. A., Jahn, A. L., & O'Shea, J. P. (2005). Toward an Objective Characterization of an Anhedonic Phenotype: A Signal-Detection Approach. *Biological Psychiatry*, *57*(4), 319-327. DOI: 10.1016/j.biopsych.2004.11.026

Pringle, A. & Harmer, C. J. (2015). The effects of drugs on human models of emotional processing: an account of antidepressant drug treatment. *Dialogues in Clinical Neuroscience*, *17*(4), 477-487.

Pulcu, E. & Browning, M. (2017). Affective bias as a rational response to the statistics of rewards and punishments. *eLife*, *6*, e27879. DOI: 10.7554/eLife.27879

Raymond, J. G., Steele, J. D., & Seriès, P. (2017). Modeling Trait Anxiety: From Computational Processes to Personality. *Frontiers in Psychiatry*, *8*, 1-19. DOI: 10.3389/fpsyt.2017.00001

Rescorla, R.A., & Wagner, A.R. (1972). A theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement. In A.H. Black & W.F. Prokasy, Eds., *Classical Conditioning II*, pp. 64–99. Appleton-Century-Crofts.

Reverdy, P., & Leonard, N. E. (2016). Parameter Estimation in Softmax Decision-Making Models With Linear Objective Functions. *IEEE Transactions on Automation Science and Engineering*, *13*(1), 54-67. DOI: 10.1109/TASE.2015.2499244

Rinck, M., & Becker, E. S. (2005). A Comparison of Attentional Biases and Memory Biases in Women with Social Phobia and Major Depression. *Journal of Abnormal Psychology*, *114*(1), 62-74. DOI: 10.1037/0021-843X.114.1.62

Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: a systematic review and meta-analysis. *Psychological Medicine*, *44*(10), 2029-2040. DOI: 10.1017/S0033291713002535

Roy-Byrne, P., Post, R. M., Uhde, T. W., Porcu, T., & Davis, D. (1985). The longitudinal course of recurrent affective illness: Life chart data from research patients at the NIMH. *Acta Psychiatrica Scandinavica*, *71*(s317), 1-33. DOI: 10.1111/j.1600-0447.1985.tb10510.x

Rude, S. S., Wenzlaff, R. M., Gibbs, B., Vane, J., & Whitney, T. (2002). Negative processing biases predict subsequent depressive symptoms. *Cognition & Emotion*, *16*(3), 423-440. DOI: 10.1080/02699930143000554

Rush, A. J., Carmody, T., & Reimitz, P. E. (2000). The Inventory of Depressive Symptomatology (IDS): Clinician (IDS-C) and Self-Report (IDS-SR) ratings of depressive symptoms. *International Journal of Methods in Psychiatric Research*, *9*(2), 45-59. DOI: 10.1002/mpr.79

Rush, A. J., Gullion, C. M., Basco, M. R., Jarrett, R. B., & Trivedi, M. H. (1996). The Inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychological Medicine*, *26*(3), 477-486. DOI: 10.1017/S0033291700035558

Rush, A. J., Trivedi, M. H., Ibrahim, H. M., Carmody, T. J., Arnow, B., Klein, D. N., ... & Thase, M. E. (2003). The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biological Psychiatry*, *54*(5), 573-583. DOI: 10.1016/S0006-3223(02)01866-8

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*(5306), 1593-1599. DOI: 10.1126/science.275.5306.1593

Segerstrom, S. C., Tsao, J. C., Alden, L. E., & Craske, M. G. (2000). Worry and Rumination: Repetitive Thought as a Concomitant and Predictor of Negative Mood. *Cognitive Therapy & Research*, *24*(6), 671-688. DOI: 10.1023/A:1005587311498

Sloan, D. M., Strauss, M. E., Quirk, S. W., & Sajatovic, M. (1997). Subjective and expressive emotional responses in depression. *Journal of Affective Disorders*, *46*(2), 135-141. DOI: 10.1016/S0165-0327(97)00097-9

Sloan, D. M., Strauss, M. E., & Wisner, K. L. (2001). Diminished Response to Pleasant Stimuli by Depressed Women. *Journal of Abnormal Psychology*, *110*(3), 488. DOI: 10.1037/0021-843X.110.3.488

Snaith, R. P., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D., & Trigwell, P. (1995). A Scale for the Assessment of Hedonic Tone: The Snaith-Hamilton

Pleasure Scale. *The British Journal of Psychiatry*, *167*(1), 99-103. DOI: 10.1192/bjp.167.1.99

Spasojević, J., & Alloy, L. B. (2001). Rumination as a Common Mechanism Relating Depressive Risk Factors to Depression. *Emotion*, *1*(1), 25. DOI: 10.1037/1528-3542.1.1.25

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1, No. 1). Cambridge: MIT press.

Thomas, C. M., & Morris, S. (2003). Cost of depression among adults in England in 2000. *The British Journal of Psychiatry*, *183*(6), 514-519. DOI: 10.1192/00-000

Tranter, R., Bell, D., Gutting, P., Harmer, C., Healy, D., & Anderson, I. M. (2009). The effect of serotonergic and noradrenergic antidepressants on face emotion processing in depressed patients. *Journal of Affective Disorders*, *118*(1), 87-93. DOI: 10.1016/j.jad.2009.01.028

Treynor, W., Gonzalez, R., & Nolen-Hoeksema, S. (2003). Rumination Reconsidered: A Psychometric Analysis. *Cognitive Therapy and Research*, *27*(3), 247-259. DOI: 10.1023/A:1023910315561

Vos, T., Barber, R. M., Bell, B., Bertozzi-Villa, A., Biryukov, S., Bolliger, I., ... & Duan, L. (2015). Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, *386*(9995), 743-800. DOI: 10.1016/S0140-6736(15)60692-4

Yu, A. J. & Dayan, P. (2005). Uncertainty, Neuromodulation, and Attention. *Neuron*, *46*(4), 681-692. DOI: 10.1016/j.neuron.2005.04.026

APPENDICES

[Appendices: DAQ]

Subject Number: _____ Date: _____

DAQ

Below is a list of statements dealing with how you generally feel about yourself and things that happen to you. Please circle the appropriate number to indicate how much you agree with each statement.

		Not At All	A Little	Somewhat	Strongly	Very Strongly
1	When bad things happen, I think it is my fault.	0	1	2	3	4
2	I feel helpless when bad things happen.	0	1	2	3	4
3	When things do not go well, I get easily discouraged.	0	1	2	3	4
4	When things go well, I think it is just due to good luck.	0	1	2	3	4
5	When something I do goes wrong, I think it is because I am incapable.	0	1	2	3	4
6	When something good happens, I think it will not last long.	0	1	2	3	4
7	When something bad happens, I think there is little I can do to make things better.	0	1	2	3	4
8	When something good happens to me, I think this was because of other people or the circumstances rather than me.	0	1	2	3	4
9	Bad things always happen to me.	0	1	2	3	4
10	When bad things happen, I rely on other people to sort things out.	0	1	2	3	4
11	When bad things happen to me, I am sure it will happen again.	0	1	2	3	4
12	When bad things happen to me, I think my life will never get better.	0	1	2	3	4
13	When something bad happens, I think of the problems this will cause in all areas of my life.	0	1	2	3	4
14	Bad things happen in all areas of my life.	0	1	2	3	4
15	When bad things happen to me, I can't see anything positive in my life.	0	1	2	3	4
16	When bad things happen, nothing seems to be in place any more.	0	1	2	3	4

[Appendices: MADRS]

STRUCTURED INTERVIEW GUIDE FOR THE

MONTGOMERY AND ASBERG DEPRESSION RATING SCALE (SIGMA)

Janet B.W. Williams, Ph.D. and Kenneth A. Kobak, Ph.D.

INTERVIEWING GUIDELINES: The questions in bold for each item should be asked exactly as written unless the information has been previously obtained, in which case it is appropriate to restate the information for confirmation. Follow-up questions are provided for use when further exploration or additional clarification of symptoms is necessary. The specified questions should be asked until you have enough information to rate the item confidently. In some cases, you may also have to add your own follow-up questions to obtain necessary information. Note that questions in parentheses are optional, for use, for example, if information is unknown. Statements in ALL CAPITALS are interviewer instructions and should not be read to the subject.

RATING GUIDELINES: Ratings should be based on the subject's condition as observed in the past week (past 7 days). As specified in the item descriptions, three of the items, Reduced Sleep, Reduced Appetite, and Inability to Feel, are rated as present only when they reflect a change from before the depression began (EUTHYMIC BASELINE). The interviewer should attempt to identify the most recent 2-month period of non-depressed functioning and use this as a reference point. In some cases, such as when the subject has dysthymia, the referent should be to the last time the subject felt alright (i.e., not depressed or high) for at least a few weeks. When a clear euthymic baseline cannot be established because of chronic depressive symptoms, these items should be rated as observed over the past 7 days instead of comparing to a previous time point.

This interview guide is based on the Montgomery-Asberg Depression Rating Scale (MADRS) (Montgomery SA, Asberg M: A new depression scale designed to be sensitive to change. *Br J Psychiatry* ; 1979 **134:** 382-9). The scale itself has been retained in its original form, except for reversing the order of the first two items. This guide adds interview questions to aid in the assessment and application of the MADRS. Previous versions of this guide appeared in 1988, 1992, 1996, 2005 and 2008.

© 2008, 2011 The Royal College of Psychiatrists. The SIGMA may be copied by individual researchers or clinicians for their own use without seeking permission from the publishers. The scale must be copied in full and all copies must acknowledge the following source: Williams JBW, Kobak KA. Development and reliability of a structured interview guide for the Montgomery-Asberg Depression Rating Scale (SIGMA). *Br J Psychiatry* 2008; **192:** 52-58. Brianne Brown, PsyD, contributed to this revision. Written permission must be obtained from the Royal College of Psychiatrists for copying, distribution to others, for replication (in print, online or by any other medium), and translations. Scientific correspondence should be addressed to Dr. Janet Williams at jbw5@columbia.edu. To inform an ongoing survey, researchers and clinicians are asked to notify Dr. Williams of their intention to use the SIGMA.

STRUCTURED INTERVIEW GUIDE FOR THE
MONTGOMERY AND ASBERG DEPRESSION RATING SCALE (SIGMA)

PT'S INITIALS: _____ PT'S ID: _____ TIME BEGAN SIGMA: _____ AM / PM

DATE: _____

INTERVIEWER: _____

OVERVIEW: **I'd like to ask you some questions about the past week. How have you been feeling since last (DAY OF WEEK)?** IF OUTPATIENT: Have you been working? (What kind of work do you do? Have you been able to work your normal hours?) IF NOT WORKING OR WORKING LESS, CLARIFY WHY.	
In the past week, have you been feeling sad or unhappy? (Depressed at all?) IF YES: Can you describe what this has been like for you? (IF UNKNOWN: How bad has that been?) IF DEPRESSED: Does the feeling lift at all if something good happens? How much does your mood lift? Does the feeling ever go away completely? (How often have you had lifts in your mood this week? What things have made you feel better?) How often did you feel (depressed/OWN EQUIVALENT) this past week? (IF UNKNOWN: How many days this week did you feel that way? How much of each day?) **In the past week, how have you been feeling about the future?** (Have you been discouraged or pessimistic?) What have your thoughts been? How (discouraged or pessimistic) have you been? How often have you felt that way? Do you think things will ever get better for you? **ESTABLISH EUTHYMIC BASELINE:** When was the last time you were well, not depressed at all, for at least 2 months?	**1. REPORTED SADNESS.** Representing reports of depressed mood, regardless of whether it is reflected in appearance or not. Includes low spirits, despondency or the feeling of being beyond help and without hope. Rate according to intensity, duration, and the extent to which the mood is reported to be influenced by events. 0 - Occasional sadness in keeping with the circumstances. 1 – 2 - Sad or low but brightens up without difficulty. 3 – 4 - Pervasive feelings of sadness or gloominess. The mood is still influenced by external circumstances. 5 – 6 – Continuous or unvarying sadness, misery, or despondency.

RATING BASED ON OBSERVATION DURING INTERVIEW AND THE FOLLOWING QUESTIONS. **In the past week, do you think you have looked sad or depressed to other people?** Did anyone say you looked sad or down? **How about when you've looked in the mirror; did you look gloomy or depressed?** IF YES: How sad or depressed do you think you have looked? How much of the time over the past week do you think you have looked depressed or down? **Has it been hard for you to laugh or smile in the past week?**	**2. APPARENT SADNESS.** Representing despondency, gloom and despair. (More than just ordinary transient low spirits) reflected in speech, facial expressions, and posture. Rate by depth and inability to brighten up. 0 – No sadness 1 – 2 - Looks dispirited but does brighten up without difficulty. 3 - 4 - Appears sad and unhappy most of the time. 5 – 6 – Looks miserable all the time. Extremely despondent.

Have you felt tense or edgy in the last week? Have you felt anxious or nervous? IF YES: Can you describe what that has been like for you? How bad has it been? **What about feeling fearful that something bad is about to happen?** How much of the time have you felt (anxious/tense/OWN EQUIVALENT) over the past week? **Have you felt panicky in the past week? IF YES:** Can you describe this feeling? How often have you felt this way? IF YES TO ANY TENSION SYMPTOM: How hard has it been to control these feelings? (What has it taken to help you feel calmer? Has anything worked to calm	**3. INNER TENSION.** Representing feelings of ill-defined discomfort, edginess, inner turmoil, mental tension mounting to either panic, dread, or anguish. Rate according to intensity, frequency, duration and the extent of reassurance called for. 0 - Placid. Only fleeting inner tension. 1 - 2 - Occasional feelings of edginess and ill-defined discomfort. 3 – 4 – Continuous feelings of inner tension or intermittent panic which the patient can only master with some difficulty. 5 – 6 – Unrelenting dread or anguish. Overwhelming panic.

you down?)	

How has your sleeping been in the last week? (How many hours have you been sleeping, compared to usual?) **Have you had trouble falling asleep?** (How long has it been taking you to fall asleep this past week? How many nights?) **Have you been able to stay asleep through the night? (**Have you been waking up at all in the middle of the night? How long does it take you to fall back to sleep? How many nights?) **Have there been any mornings this past week when you have awakened earlier than (EUTHYMIC BASELINE)?** IF UNKNOWN: Has your sleeping been restless or disturbed?	**4. REDUCED SLEEP.** Representing the experience of reduced duration or depth of sleep compared to the subject's own normal pattern when well. 0 - Sleeps as usual. 1 - 2 - Slight difficulty dropping off to sleep or slightly reduced, light, or fitful sleep. 3 – 4 – Sleep reduced or broken by at least 2 hours. 5 – 6 – Less than 2 or 3 hours sleep.

How has your appetite been this past week? (What about compared to your usual appetite?) IF NOT REDUCED: Have you been less interested in food? (How much less?) **Does food taste as good as usual?** IF LESS: How much less? Does it have any taste at all? (Have you had to push yourself to eat or have other people had to urge you to eat?)	**5. REDUCED APPETITE.** Representing the feeling of a loss of appetite compared with when well. Rate by loss of desire for food or the need to force oneself to eat. 0 - Normal or increased appetite. 1 – 2 - Slightly reduced appetite. 3 – 4 – No appetite. Food is tasteless. 5 – 6 – Needs persuasion to eat at all.

Have you had trouble concentrating or collecting your thoughts in the past week? (How about at home or at work?) IF YES: Can you give me some examples? (Have you been able to concentrate on reading a book or on the computer? Do you need to read things over and over again? Are you able to follow movies or television?) How often has that happened in the past week? Has this caused any problems for you? **Have you had any trouble following a conversation?** (IF YES: How bad has that been? How often has that happened this past week?) **NOTE:** ALSO CONSIDER BEHAVIOR DURING INTERVIEW.	**6. CONCENTRATION DIFFICULTIES.** Representing difficulties in collecting one's thoughts mounting to incapacitating lack of concentration. Rate according to intensity, frequency, and degree of incapacity produced. 0 - No difficulties in concentration. 1 - 2 - Occasional difficulties in collecting one's thoughts. 3 – 4 – Difficulties in concentrating and sustaining thought which reduces ability to read or hold a conversation. 5 – 6 – Unable to read or converse without great difficulty.

Have you had any trouble getting started at things in the past week? IF YES: What things? How bad has that been? Have you had difficulty getting started at simple routine everyday things (like getting dressed, brushing your teeth, showering)? Are you OK once you get started at things or is it still more of an effort to get something done? Has there been anything that you needed to do that you were unable to do? Have you needed help to do things? IF YES: What things? How often? **Have you done everyday things more slowly than usual? IF YES: Like what, for example? How bad has that been?**	**7. LASSITUDE.** Representing a difficulty getting started, or slowness initiating and performing everyday activities. 0 - Hardly any difficulty in getting started. No sluggishness. 1 - 2 - Difficulties in starting activities. 3 – 4 – Difficulties in simple routine activities, which are carried out with effort. 5 – 6 – Complete lassitude. Unable to do anything without help.

Have you been less interested in things around you, or in activities you used to enjoy? IF YES: What things? How much less interested in (those things) are you now compared to (EUTHYMIC BASELINE)? **What things have you enjoyed this week?** How much did you enjoy them? **Has there been any change in your ability to feel emotions?** (Do you feel things less intensely than you used to, things like anger, grief, pleasure?) IF YES: Can you tell me more about that? (IF UNKNOWN: Are you able to feel any emotions at all?) **Have your feelings towards family and friends changed at all?** IF YES: Do you feel less towards them than you used to?	**8. INABILITY TO FEEL.** Representing the subjective experience of reduced interest in the surroundings, or activities that normally give pleasure. The ability to react with adequate emotion to circumstances or people is reduced. 0 - Normal interest in the surroundings and in other people. 1 - 2 - Reduced ability to enjoy usual interests. 3 – 4 – Loss of interest in the surroundings. Loss of feelings for friends and acquaintances. 5 – 6 – The experience of being emotionally paralyzed, inability to feel anger, grief or pleasure, and a complete or even painful failure to feel for close relatives and friends.

Have you been putting yourself down, or feeling that you're a failure in some way, over the past week? (Have you been blaming yourself for things that you've done, or not done?) IF YES: What have your thoughts been? How often have you felt that way? **In the past week have you been feeling guilty about anything? What about feeling as if you have done something bad or sinful?** IF YES: What have your thoughts been? How often have you felt that way? ALSO CONSIDER RESPONSES TO QUESTIONS ABOUT PESSIMISM FROM ITEM 1.	**9. PESSIMISTIC THOUGHTS.** Representing thoughts of guilt, inferiority, self-reproach, sinfulness, remorse, and ruin. 0 - No pessimistic thoughts. 1 - 2 - Fluctuating ideas of failure, self-reproach, or self-depreciation. 3 – 4 – Persistent self-accusations, or definite but still rational ideas of guilt or sin. Increasingly pessimistic about the future. 5 – 6 – Delusions of ruin, remorse, or unredeemable sin. Self-accusations which are absurd and unshakeable.

This past week, have you felt like life isn't worth living? (IF NO: What about feeling as if you're tired of living?) IF YES: Tell me about that. How often have you felt that way? **This week, have you thought that you would be better off dead?** IF YES: Tell me about that. How often have you felt that way? Have you had thoughts of hurting or even killing yourself this past week? **IF YES: What have you thought about? How often have you had these thoughts? How long have they lasted? Have you actually made plans? IF YES: What are these plans? Have you made any preparations to carry out these plans? (Have you told anyone about it?)**	**10. SUICIDAL THOUGHTS.** Representing the feeling that life is not worth living, that a natural death would be welcome, suicidal thoughts, and preparation for suicide. Suicidal attempts should not in themselves influence this rating. 0 - Enjoys life or takes it as it comes. 1 – 2 - Weary of life. Only fleeting suicidal thoughts. 3 - 4 - Probably better off dead. Suicidal thoughts are common, and suicide is considered as a possible solution, but without specific plans or intention. 5 – 6 – Explicit plans for suicide when there is an opportunity. Active preparations for suicide.

TIME ENDED SIGMA:	**_____ AM / PM**
TOTAL MADRS SCALE SCORE:	**_____ _____**

[Appendices: RSS]

Subject Number: _____ Date: _____

Ruminative Response Scale (RSS)

People think and do many different things when they feel depressed. Please read each of the items below and indicate whether you almost never, sometimes, often, or almost always think or do each one when you feel down, sad, or depressed. Please indicate what you generally do, not what you think you should do.

[Points]

1: Almost never 2: Sometimes 3: Often 4: Almost always

1. Think about how alone you feel
2. Think “I won’t be able to do my job if I don’t snap out of this.”
3. Think about your feelings of fatigue and achiness
4. Think about how hard it is to concentrate
5. Think “What am I doing to deserve this?”
6. Think about how passive and unmotivated you feel
7. Analyse recent events to try to understand why you are depressed
8. Think about how you don’t seem to feel anything anymore
9. Think “Why can’t I get going?”
10. Think “Why do I always react this way?”
11. Go away by yourself and think about why you feel this way
12. Write down what you are thinking about and analyse it
13. Think about a recent situation, wishing it had gone better
14. Think “I won’t be able to concentrate if I keep feeling this way.”
15. Think “Why do I have problems other people don’t have?”
16. Think “Why can’t I handle things better?”
17. Think about how sad you feel
18. Think about all your shortcomings, failings, faults and mistakes
19. Think about how you don’t feel up to doing anything
20. Analyse your personality to try to understand why you are depressed
21. Go someplace alone to think about your feelings
22. Think about how angry you are with yourself

[Appendices: SHAPS]

Subject Number: _____ Date: _____

Snaith–Hamilton Pleasure Scale (SHAPS)

This questionnaire is designed to measure your ability to experience pleasure in the last few days. It is important to read each statement very carefully. Tick one of the boxes to indicate how much you agree or disagree with each statement.

I would enjoy my favourite television programme	Strongly Disagree	Disagree	Agree	Strongly Agree
I would enjoy being with my family or close friends	Definitely Agree	Agree	Disagree	Strongly Disagree
I would find pleasure in my hobbies and pastimes	Strongly Disagree	Disagree	Agree	Strongly Agree
I would be able to enjoy my favourite meal	Definitely Agree	Agree	Disagree	Strongly Disagree
I would enjoy a warm bath or refreshing shower	Definitely Agree	Agree	Disagree	Strongly Disagree
I would find pleasure in the scent of flowers or the smell of a fresh sea breeze or freshly baked bread	Strongly Disagree	Disagree	Agree	Strongly Agree
I would enjoy seeing other people's smiling faces	Definitely Agree	Agree	Disagree	Strongly Disagree
I would enjoy looking smart when I have made an effort with my appearance	Strongly Disagree	Disagree	Agree	Strongly Agree
I would enjoy reading a book, magazine or newspaper	Definitely Agree	Agree	Disagree	Strongly Disagree
I would enjoy a cup of tea or coffee or my favourite drink	Strongly Disagree	Disagree	Agree	Strongly Agree
I would find pleasure in small things, e.g. bright sunny day, a telephone call from a friend	Strongly Disagree	Disagree	Agree	Strongly Agree
I would be able to enjoy a beautiful landscape or view	Definitely Agree	Agree	Disagree	Strongly Disagree
I would get pleasure from helping others	Strongly Disagree	Disagree	Agree	Strongly Agree
I would feel pleasure when I receive praise from other people	Definitely Agree	Agree	Disagree	Strongly Disagree

[Appendices: Trait-STAI]

Subject Number: _____ Date: _____

State-Trait Anxiety Inventory (STAI): Trait Questionnaire

A number of statements which people have used to describe themselves are given below. Read each statement and then circle the appropriate number to the right of the statement to indicate how you ***generally*** feel. There are no right and wrong answers. Do not spend too much time on each statement but give the answer which seems to describe you best.

	Not at all	Somewhat	Moderately	Very much
I feel pleasant.	1	2	3	4
I feel nervous and restless.	1	2	3	4
I am satisfied with myself.	1	2	3	4
I wish I could be as happy as others seem to be.	1	2	3	4
I feel like a failure.	1	2	3	4
I feel rested.	1	2	3	4
I am “cool, calm and collected.”	1	2	3	4
I feel that the difficulties are piling up so that I cannot overcome them.	1	2	3	4
I worry too much over something that does not really matter.	1	2	3	4
I am happy.	1	2	3	4
I have disturbing thoughts.	1	2	3	4
I lack self-confidence.	1	2	3	4
I feel secure.	1	2	3	4
I make decisions easily.	1	2	3	4
I feel inadequate.	1	2	3	4
I am content.	1	2	3	4
Some unimportant thoughts run through my mind and bother me.	1	2	3	4
I take disappointments so keenly that I cannot put them out of my mind.	1	2	3	4
I am a steady person.	1	2	3	4
I get in a state of tension or turmoil as I think over recent concerns and interests.	1	2	3	4