Newcastle University

NEUROPSYCHOLOGICAL PERFORMANCE, EMOTION PROCESSING AND PSYCHOSOCIAL FUNCTION IN BIPOLAR DISORDER

Lucy Jayne Robinson

Thesis submitted in partial fulfillment of the requirements
of the regulations for the degree of Doctor of Philosophy

Newcastle University

Faculty of Medical Sciences

Institute of Neuroscience

June 2010

CONTENTS

Contents 2

Acknowledgements 9

Abstract 11

Chapter 1: Bipolar disorder and its diagnosis 13

The disorder 13

Overview 13

The bipolar spectrum 13

Symptoms and presentation 14

Depression 14

Mania and hypomania 15

Mixed states 17

Cycle frequency 18

Incidence and prevalence 18

Age at onset 19

Gender 19

Rapid Cycling 20

Aetiology 20

Genetics 20

Neurohormonal Abnormalities 22

Hypothalamic-Pituitary-Adrenal (HPA) axis dysfunction 22

Hypothalamic-Pituitary-Thyroid (HPT) axis and rapid cycling 23

Structural brain differences 23

Psychosocial influences 24

Comorbidity 26

Course and prognosis 26

Neuropsychological function 27

Chapter 2: Introduction to Neuropsychological Function in Bipolar Disorder 30
Cognitive Function in Euthymia 31
Cognitive Dysfunction as an Endophenotype 33
Aetiological issues 37
Longitudinal studies 37
Prospective study 37
Retrospective studies 37
Post-onset longitudinal studies 39
Cross-sectional studies 41
Single versus multiple episodes 41
Relationship with illness features 42
Unaffected First Degree Relatives 43
Summary 43
Key questions 43
Relationship between executive and memory dysfunction 43
Is cognitive impairment in bipolar disorder modifiable? 45
Relationship with Social Functioning 45
Chapter 3: Introduction to Emotion Processing in bipolar disorder 47
Introduction 47
Mood 47
Emotion 50
Models of emotion 51
Emotion, Cognition and affective disorder 55
Deficit or bias 55
Explicit measures 56
Implicit measures 59
Psychology of bipolar disorder and the Manic defence 61
Explicit Emotion recognition in bipolar disorder 63
Facial Expression Recognition 63

Manic patients 63
Depressed patients 64
Euthymic patients 65
Control tasks 65
Vocal Emotion Recognition 66
Secondary emotions 66
Implicit Emotion Tasks 70
Summary and Issues in Bipolar Disorder 73
Chapter 4: Introduction to Psychosocial Functioning in Bipolar Disorder 76
What is "Function"? 76
Measurement of Functioning 79
Functioning in bipolar Disorder 82
Predictors of functioning 83
Cognition and functioning 85
Discussion 89
Chapter 5: Neuropsychological function in Bipolar Disorder 94
Introduction 94
Methods 97
Participants - Patients with Bipolar Disorder 97
Participants – Controls 99
Ethics 100
Measures – Self-rated Questionnaires 100
Measures – Neuropsychological Tests 102
Procedure 114
Data Analysis 115
Results 117
Demographics 117
Neuropsychological function 120
Summary of results 127

Discussion 128

Chapter 6: Memory and Executive Function in Bipolar Disorder 131

Introduction 131

Methods 141

Participants 141

Materials 143

Procedure 144

Data Analysis 145

Results 148

Depressed Patients – Standard Administration: 148

Euthymic Patients – Standard Administration: 153

Euthymic Patients – Shuffled Administration: 159

Exploratory Analysis of the Relationship between Subjective Organisation Measures and Executive Function 164

Relationship Between Subjective Organisation Indices 168

Discussion 168

Chapter 7: Modifying Cognitive Impairment in Bipolar Disorder 177

Introduction 177

Methods 180

Participants 180

Materials 180

Procedure 182

Data Analysis 182

Results 183

Discussion 187

Chapter 8: Emotion Processing in Bipolar Disorder 192

Introduction 192

Facial expression recognition 192

Vocal emotion recognition 193

Implicit Emotion Processing 193
Methods 195
Participants 195
Materials 195
Procedure 207
Data Analysis 208
Results 208
Facial Expression Recognition Task 208
Vocal Expression Recognition Task 213
Facial Dot Probe 214
Semantic Stroop 215
Reading the Mind in the Eyes Test 217
Discussion 217
Chapter 9: Psychosocial Functioning in Bipolar Disorder 225
Introduction 225
Methods 226
Participants 226
Materials 226
Procedure 231
Results 241
Clinician-Rated measures 241
Participant-rated measures 243
Predicting Functioning 245
Patients - General 245
Patients - Clinican-rated vs self-rated measures 246
Patients - Differences between functional domains 247
Patients - The role of demographic factors and clinical variables 249
Controls - General 251
Controls - Clinican-rated vs self-rated measures 252

Controls - Differences between functional domains 252
Controls - The role of demographic factors 252
Summary 255
Factor Analysis 255
Cognition 255
Emotion-Processing 257
Psychological Measures 258
Final Overall Model of Functioning 259
Learning Potential and Functioning 263
Discussion 264
Summary – Between-group Results 264
Summary - Predicting Functioning 264
Interpretation and Relationship with Previous Findings 266
Limitations 271
Conclusion & future directions 275
Chapter 10: Discussion 279
Summary of Findings 279
Discussion 281
Limitations 283
Future Directions 284
References 287
Appendices 310
Appendix 1: Literature search criteria 310
Appendix 2: Flow chart of Participants 311
Appendix 3: Study Procedure and Order of Administration 312
Appendix 4: Neuropsychological impairment Covarying for IQ and Education 314
Appendix 5: Recall & Subjective Organisation Graphs 317
Appendix 6: Correlation Matrices 320
Appendix 7: Psychosocial Functioning Scales 324

ACKNOWLEDGEMENTS

I would firstly like to thank the volunteers who gave their time and energy to come and take part. For some it was more difficult and more stressful than for others and I greatly appreciate your help and the spirit with which you participated – to help further our knowledge of this illness in the hope that, in the longer term, we can gradually put the pieces together and come to understand this puzzle a lot better.

Secondly, immense thanks go to my supervisors, both of whom offered a unique contribution along this journey. Wise words and tolerance were rarely far away. Thank you Nicol for your pragmatism and support, especially for providing well-needed boosts at just the right times. Thank you John for your keen critical appraisal that did its best to ensure I was not left perching dangerously on any laurels for too long. I have learned important lessons from you both that I will take forwards with me.

Especial thanks to Peter Gallagher for seemingly endless support, patience, kindness and motivation along a road that got bumpy at times. Thanks for riding it out with me.

A large number of family, friends and colleagues also deserve acknowledgement for their encouragement, their willingness to act as guinea pigs and pilot test subjects, and for their ready supply of distraction and entertainment that made any procrastination far more fun than it would have been otherwise. Thanks to Mum for being my first test participant – oh and also for my monthly cakes and loaning me the dog when needed. Thanks to Alex for my daily digests of vital and sustaining information, and for being an all-round excellent brother and running comrade. Thanks to Gill, another endurance sports nut, for testing out some of the tasks for me and for sharing all those arduous, stress-relieving swim sessions. Thanks to many friends, but especially Jenny and Michaela for keeping me looking forwards, for having lots of fun, and for reminding me that any progress, however small, is progress

nonetheless. Thanks to Melanie for your kind and gracious supervision that frequently strayed into PhD-related matters. Your experiences and patience were greatly appreciated. Finally thanks to all the members of North Shields Poly Athletics Club who have kept me going in so many different ways over the last three years.

ABSTRACT

Background/aims: There is considerable evidence that euthymic individuals with bipolar disorder experience both significant impairment in cognitive function – especially verbal memory and executive function – and significant impairment in psychosocial function. There is also some evidence that bipolar disorder is associated with altered emotional processing, although the findings are somewhat mixed. The aims of the present study were to address some of the outstanding questions about these three areas in euthymic bipolar patients: 1) is there a relationship between verbal memory impairment and executive dysfunction? 2) Can cognitive function be improved? 3) Do patients show differences in the explicit and/or implicit processing of emotional information? 4) Is there a relationship between cognitive dysfunction, emotion-processing impairment, and psychosocial functioning?

Method: A large sample of euthymic patients with bipolar disorder and healthy controls underwent a battery of neuropsychological tests, tests of emotion processing, and had their social function measured via a selection of clinician- and self-rated measures.

Results: 1) Deficits in executive function were insufficient to explain the entire extent of the verbal memory impairment in this group, suggesting they are two distinct areas of impairment. 2) A subgroup of patients who underwent a simple self-monitoring intervention significantly outperformed a control group of patients on several indices of one cognitive functioning measure. 3) Patients showed little evidence of impairment in emotion processing on either explicit or implicit measures. 4) Functioning was most strongly predicted by psychological factors, such as self-esteem, anxiety and dysfunctional attitudes. Cognitive function was also a significant, although weaker, predictor. Overall models accounted for more than 60% of the variance in functioning.

Conclusion: Cognitive dysfunction in bipolar disorder appears to be multi-faceted and involve more than one area of impairment. Efforts to improve functional outcomes would be best served focusing on improving self-esteem and psychological coping resources.

CHAPTER 1: BIPOLAR DISORDER AND ITS DIAGNOSIS

THE DISORDER

OVERVIEW

Bipolar disorder is a cyclical mood disorder involving periods of profound disruption to mood and behaviour interspersed with periods of more or less full recovery. Most commonly sufferers experience both episodes of low mood (depression) and elevated mood (mania or hypomania). The diagnosis, originally termed manic depressive insanity, was postulated by the psychiatrist Emil Kraepelin around the end of the 19th century, although the historical record contains description of the frenetic activity associated with the manic state dating back to the ancient Greeks and even the ancient Egyptians. In 1980 the name bipolar disorder was adopted to replace the older term. In this modern conceptualisation the key feature of bipolar disorder is the experience of hypomania or mania – grandiose and expansive affect associated with increased drive and decreased sleep, which ultimately can culminate in psychosis and exhaustion if left untreated.

THE BIPOLAR SPECTRUM

Far from being a discrete diagnostic entity, there is increasing recognition of a spectrum of bipolar disorders that range from marked and severe mood disturbance into milder mood variations that are much closer to non-pathological levels of mood fluctuation. In terms of classification, in the Diagnostic and Statistical Manual 4th Edition (DSM-IV; APA, 1994) a distinction is drawn between bipolar I disorder, in which the patient suffers full-blown manic episodes (most commonly interspersed with episodes of major depression), and bipolar II disorder, in which the patient experiences depressive episodes and hypomanic

episodes. Cyclothymia, another bipolar disorder described in DSM-IV, involves recurrent hypomanic episodes and subclinical episodes of depression. 'Softer' forms of bipolar disorder have been proposed (Akiskal et al., 2000). However, owing to difficulty establishing both the clinical utility of these potential diagnoses and satisfactory inter-rater reliability in their assessment, they are not currently part of official diagnostic classifications. In fact, the considerable phenotypic variation in the presentation of bipolar disorder is becoming increasingly problematic for research into the condition. It is increasingly common for research studies to include analyses of subgroups of patients with specific symptom clusters or illness features (e.g. anhedonia, psychosis, early onset) in order to better understand the sources of variation between patients in both clinical picture and response to treatment. The present investigation focuses only on patients with bipolar I or bipolar II disorder.

SYMPTOMS AND PRESENTATION

DEPRESSION

Although mania or hypomania are the defining characteristics of bipolar disorder, throughout the course of the illness depressive symptoms are more common than manic symptoms. Patients with bipolar disorder spend a substantial proportion of time suffering from syndromal or subsyndromal depressive symptoms. The outcome of a 12-year prospective longitudinal study, in which 146 patients with bipolar I disorder completed weekly mood ratings, reported that depressive symptoms were three times more common than manic or hypomanic symptoms (Judd et al., 2002). Patients spent 32% of weeks with symptoms of depression. In a separate study of 86 patients with bipolar II disorder this proportion was much higher at 50% (Judd et al., 2003). A similar study by the Stanley Bipolar Network monitored 258 bipolar patients for a year and reported that on average patients spent 33% of the time depressed and a large proportion (>60%) suffered four or more mood episodes in a year (Post et al., 2003).

Major depressive episodes in bipolar disorder are similar to those experienced in unipolar major depression. Patients suffer depressed mood and experience profound loss of interest in activities coupled with other symptoms such as fatigue, weight loss or gain, insomnia or hypersomnia, psychomotor slowing, feelings of worthlessness, excessive guilt, and suicidal thoughts or actions. However, evidence suggests that depression in the course of bipolar disorder may be more likely to show signs of psychomotor retardation, to have melancholic features (such as feelings of worthlessness and marked anhedonia), to show features of atypical depression (such as hypersomnia and weight gain) (Mitchell and Malhi, 2004), and to show psychotic features – especially in young people (Strober and Carlson, 1982). Subsyndromal depressive symptoms are common in patients with bipolar disorder (especially those with bipolar II disorder) and are often associated with significant interpersonal or occupational disability (Morriss, 2002).

The risk of suicide is greatly elevated during depressive episodes. Approximately 17% of patients with bipolar I disorder and 24% of patients with bipolar II disorder attempt suicide during the course of their illness (Rihmer and Kiss, 2002). Most suicide attempts and most completed suicides occur in the depressed phase of the illness and patients with bipolar II disorder are at especially high risk (Baldessarini et al., 2003).

MANIA AND HYPOMANIA

The longitudinal study of bipolar symptomatology mentioned above reported that patients with bipolar I disorder suffered syndromal or subsyndromal manic or hypomanic symptoms approximately 9% of the time over 12 years (Judd et al., 2002). For patients with bipolar II disorder approximately 1% of weeks were spent hypomanic (Judd et al., 2003). Similarly, the one-year prospective follow-up study conducted by the Stanley Bipolar Network reported that on average patients experienced syndromal manic symptoms approximately 10% of the time (Post et al., 2003). In the majority of cases individuals with

bipolar disorder will experience both manic and depressive episodes throughout the course of their illness, although one epidemiological survey identified a subpopulation of approximately 20% who had never experienced a depressive episode (Kessler et al., 1997).

Although mania is much less common than depression in those with bipolar disorder, the extreme behaviours associated with it can be devastating. Patients in the acute manic phase exhibit expansive, grandiose affect, which may be predominately euphoric or irritable. Although dysphoric mood is more frequently associated with depressive episodes, factor analytic studies of symptoms in patients with pure mania suggest dysphoric mood (such as depression, guilt and anxiety) can be prominent in some manic patients (Cassidy et al., 1998) (Cassidy and Carroll, 2001).

The clinical presentation of mania is marked by several features, which may include inflated self-esteem, disinhibition, decreased need for sleep, excessive talkativeness with pressure of speech, racing thoughts, distractibility, increased activity levels, and engaging in pleasurable activities that involve large risks (such as sexual activities or excessive spending). The early stages of mania can involve increased productivity which may feel satisfying and rewarding. However, as the episode worsens severe distractibility, restlessness, and difficulty concentrating can render the completion of tasks impossible. Sleep deprivation can result in physical exhaustion with no desire to rest. The person may find it hard to stay still or remain seated. In severe cases individuals may develop psychotic symptoms such as grandiose delusions and mood-congruent hallucinations – for example, the voice of God sending messages of special purpose. Alternatively persecutory delusions may develop, but are usually consistent with a general grandiose theme such as the belief that others are actively trying to thwart the person's plans or remove their power. Insight is lost in mania – the individual is unaware that their behaviour is abnormal and does not consider him or herself to be in need of treatment. Clinical interventions may be seen as attempts to

undermine the person's esteem and power and could provoke or worsen irritability even in patients who are predominantly euphoric.

All the features reported in mania – except psychotic symptoms – can also occur in hypomania to a less severe extent. Generally insight is preserved, although the person may not feel in need of help. Increased productivity and decreased need for sleep can be experienced as a positive enhancement of everyday functioning. Hypomania is accompanied by a change in functioning that is not characteristic of the person when non-depressed and the change is noticed by others, but it is not associated with marked impairment in social or occupational function. According to the DSM-IV diagnostic criteria, symptoms must last at least 4 days to merit the diagnosis of a hypomanic episode. However, there is considerable debate about how long hypomanic symptoms should be present to merit a diagnosis of bipolar II disorder.

MIXED STATES

In a mixed episode criteria are met for a depressive episode and a manic episode simultaneously. The patients with bipolar I disorder who took part in the 12-year longitudinal study mentioned previously spent an average 6% of weeks in a mixed or cycling state (where polarity of episode was changing and symptoms of both were present) (Judd et al., 2002). For patients with bipolar II disorder the proportion was just over 2% (Judd et al., 2003). It is estimated that approximately two thirds of patients will suffer a mixed episode at some point in their illness (Mackin and Young, 2005). A study of 441 patients with bipolar disorder reported that subclinical mixed episodes are common – with 70% of those in a depressed episode showing clinically significant signs of hypomania, and 94% of those with mania or hypomania showing significant depressive symptoms (Bauer et al., 2005). Sub-threshold mixed episodes were more than twice as prevalent as threshold mixed episodes.

The combination of morbid, depressed affect with overactivity and racing thoughts makes mixed states a particularly dangerous time for people with bipolar disorder.

CYCLE FREQUENCY

There is a large amount of variation in how often patients suffer mood episodes and no criteria exist to define 'normal' cycle frequency. Some patients have discrete episodes that occur rarely (for example, no more than one episode per year) with full recovery in between, others experience episodes more often, and some may fail to fully recover between episodes. A subset of patients suffers from rapid cycling bipolar disorder, which is defined as the experience of at least four syndromal depressive, manic, hypomanic, or mixed episodes within a twelve month period. A review of the last 30 years of research on rapid versus non-rapid cycling indicated differences in illness course and prognosis (Mackin and Young, 2004) and reports suggest the distinction is of value as a course modifier (Maj et al., 1994, Maj et al., 1999). A recent study in a sample of 456 bipolar probands identified 91 (20%) patients with a rapid cycling illness, who, in comparison to those without rapid cycling, suffered more severe mood symptoms and a greater degree of functional impairment (Schneck et al., 2004).

INCIDENCE AND PREVALENCE

Community-based epidemiological studies consistently report the lifetime prevalence of bipolar I disorder to be approximately 1% (based on DSM-III-R, or DSM-IV criteria). A review of epidemiological surveys in six non-European countries reported that the lifetime rate of bipolar I disorder ranged from 0.3% to 1.5% (Weissman et al., 1996). Rates reported in European studies have varied more widely from 0.1% to 2.4% (Faravelli et al., 1990, Pini et al., 2005, Regeer et al., 2004, Szadoczky et al., 1998, ten Have et al., 2002). Estimates of the lifetime prevalence of bipolar II disorder vary more widely due to

differences in diagnostic practices both over time and geography. One early American study estimated the lifetime risk of bipolar II disorder to be approximately 0.6% (Weissman and Myers, 1978). European studies have indicated prevalence estimates of between 0.2%-2.0% (Faravelli et al., 1990, Szadoczky et al., 1998).

AGE AT ONSET

In the review of epidemiological surveys mentioned previously, the mean age at onset reported by each of the six studies ranged from 17.1-29 years, with a peak in onset rate occurring between the ages of 15 and 19 years (Weissman et al., 1996). A large retrospective study of patients with bipolar disorder reported that there was an average 8 years' delay from a patient's first recollected mood episode to receiving a diagnosis of bipolar disorder (Mantere et al., 2004). There are no reliable prodromal signs or symptoms of bipolar disorder. One study provided some evidence of prodromal mood disturbance in patients who went on to develop bipolar disorder, however could not distinguish between patients who went on to develop a different psychiatric disorder (Thompson et al., 2003). While most episodes of bipolar disorder first present by 30 years of age, bipolar disorder can present later in life. Late onset bipolar disorder is characterised by a lower burden of family history of psychiatric disorder, greater medical comorbidities and a greater incidence of subsequent neurological problems (Fujikawa et al., 1995, Shulman, 1997). Late onset bipolar disorder may also show a greater latency between the initial depressive episode and subsequent manic episode.

GENDER

Bipolar I disorder occurs approximately equally in both sexes (Lloyd et al., 2005). The symptom profile may differ between men and women; there is some evidence that women tend to experience more episodes of mixed or dysphoric mania than men (Arnold et al., 2000). There is disputed evidence that bipolar II disorder is more common in females than

males. Recent data from a large sample of patients with bipolar disorder found a significantly higher incidence of bipolar II disorder in women (29.0%) than men (15.3%) (Baldassano et al., 2005). In a general population survey using DSM-III-R criteria (which require a minimum of 4 days of hypomanic symptoms for a hypomanic episode) there was no reported gender difference in the prevalence of bipolar II disorder (Szadoczky et al., 1998). However, a population study using broader criteria for bipolar II disorder not requiring this minimum duration found a 1-year prevalence rate for hypomania of 7.4% in females and only 2.7% in males (Angst et al., 2003).

RAPID CYCLING

Estimates of the prevalence of rapid cycling within bipolar I and II disorders have ranged from 13-56% depending on the definition of rapid cycling used, but most studies suggest approximately 20% (Mackin and Young, 2004).

AETIOLOGY

Despite its long history, little is known about what causes bipolar disorder. Recent research efforts have concentrated on identifying possible biological underpinnings of the disorder including genetic components, neurohormonal abnormalities and structural brain differences. Explanations in terms of psychosocial factors were mainstream in the 19th century, but more recently have largely given way to biological theories. However, as of yet there is no overarching explanation and the heterogeneous clinical presentation of bipolar disorder suggests that a number of different mechanisms are involved.

GENETICS

Family, twin and adoption studies conducted over the last few decades have identified that bipolar disorder has a heritable component. Studies report that first degree

relatives of an individual with bipolar disorder face a lifetime risk of developing the illness 5-10 times greater than the general population (Craddock and Jones, 2001). Studies in monozygotic and dizygotic twins where at least one twin is affected by bipolar disorder have provided support for genetic transmission. Monozygotic co-twins of bipolar probands face a 40-70% risk of developing bipolar disorder and the concordance rate of approximately 60% is markedly higher than that for dizygotic twins (Craddock and Jones, 1999). The largest twin study investigating heritability to date reported that 85% of the variance in the diagnosis of bipolar disorder was accounted for by genetic factors, with non-shared environmental influences accounting for the remaining 15% of variance (McGuffin et al., 2003).

However, the inheritance pattern is not simple and is not consistent with a single gene model of bipolar disorder. Mathematical modeling of the genetics data has indicated that the inheritance pattern is most consistent with multiple susceptibility genes, with each individual gene exerting only a small or moderate effect (Craddock and Sklar, 2009). Genetics studies using a variety of techniques have identified a number of different loci or genes as candidates for further investigation. However, in general these studies have generated more differences than similarities and very few results have been replicated. As such, no reliable genetic markers have been identified.

However, the utility of genetic studies for enhancing our understanding of the underlying causes of psychiatric illness is potentially highly limited by the degree of phenotypic variation within the disorder and also by the overlap between psychiatric illnesses. The various iterations of the DSM have essentially enshrined descriptions of the observed behavior of individuals with mental illness that have their historical roots in the late 1800s. The function of such a system is to increase the reliability with which individuals with similar symptom-clusters can be described with a common term and thus enhance clinician decision-making. This system was not designed to accurately reflect (or even

identify) the underlying biological differences or similarities behind the conditions. Naturally, to maximise reliability, categorical diagnostic systems focus more on the (outward) differences between conditions than on their similarities. The clean lines drawn by systems like the DSM have not been reflected in the inheritance studies or genetics data so far (Ivleva et al., 2010) and it may be that the commonalities between diagnoses are of greater relevance from a biological standpoint than the differences. This suggests that future genetics research needs to widen its focus and not rely on the lines currently drawn in the sand to truly reflect the strata lying below. Whether that means moving towards a dimensional approach, or focusing on specific symptoms irrespective of diagnosis requires further investigation.

NEUROHORMONAL ABNORMALITIES

Much attention recently has focussed on the role of the endocrine system in mood disorders. Interest has centred on two biological systems: the hypothalamic-pituitary-adrenal (HPA) axis, one of the major hormonal systems activated during stress, and the hypothalamic-pituitary-thyroid (HPT) axis.

HYPOTHALAMIC-PITUITARY-ADRENAL (HPA) AXIS DYSFUNCTION

In response to stress, neurons in the hypothalamus secrete the chemical messenger corticotrophin releasing hormone (CRH) to the anterior pituitary gland to stimulate the production of adrenocorticotropic hormone (ACTH) which in turn stimulates the adrenal glands to produce coritsol. Cortisol influences immune system function, has a potent anti-inflammatory action and is a major regulator of the physiological stress response. Importantly, it provides negative feedback to the hypothalamus which shuts down the stress response and eventually returns cortisol to normal, pre-stress levels. A number of studies have reported abnormalities in this system in patients with bipolar disorder which are

consistent with reduced HPA axis feedback (Rybakowski and Twardowska, 1999, Schmider et al., 1995, Watson et al., 2004). Chronically elevated levels of cortisol can have deleterious consequences, including effects on mood and memory. Interestingly, signs of HPA axis dysfunction have been observed in all stages of bipolar illness, including during remission. Such dysfunction could underlie susceptibility to future episodes and account for the often chronic course of bipolar disorder.

HYPOTHALAMIC-PITUITARY-THYROID (HPT) AXIS AND RAPID CYCLING

The HPT axis is also of interest in bipolar disorder, particularly in the genesis of rapid cycling. Abnormalities of thyroid function are noted in patients with depression and mania. Subclinical hypothyroidism is seen in a significant proportion of patients with treatment resistant depression. Along with evidence of mild hypothyroidism, patients in the manic state may show reduced responsiveness of the pituitary gland to the chemical messenger thyrotropin-releasing hormone which stimulates activity of the thyroid gland. Approximately 25% of patients with rapid cycling bipolar disorder have evidence of hypothyroidism, which contrasts with only 2-5% of depressed patients in general (Muller-Oerlinghausen et al., 2002). Since thyroid hormones have profound effects on mood and behaviour, dysfunction in the HPT axis could explain some of the presenting symptoms of patients with bipolar disorder.

STRUCTURAL BRAIN DIFFERENCES

In comparison to work in schizophrenia, there have been relatively few studies investigating structural brain differences in patients with bipolar disorder and findings have been contradictory. A major review of neuroanatomical studies in bipolar disorder reported some evidence of enlarged ventricles and abnormalities in the frontal and temporal lobes in at least a subpopulation of patients (Bearden et al., 2001). An excess of white matter lesions

has also been reported (Altshuler et al., 1995) and one study reported that the number of white matter lesions correlated negatively with functional outcome (Moore et al., 2001). Although the hope in identifying a neuroanatomical profile of bipolar disorder is to develop an understanding of neurodevelopmental or genetic contributions to the illness, it is currently unknown whether differences are the cause or consequence of affective disorder.

PSYCHOSOCIAL INFLUENCES

Although much recent research has focussed on biological factors, a number of psychosocial factors have also been identified that may be relevant to understanding the development and progression of bipolar disorder or a particular individual's presentation. Antecedent factors, such as childhood maltreatment, may act as predisposing factors for developing the disorder, whereas concurrent factors such as social class, social support, and self-esteem may act as course modifiers or precipitants for episodes (Alloy et al., 2006, Johnson and Fingerhut, 2006, Leff, 2001, Post et al., 2001).

A potential role for psychosocial stressors in both the aetiology and exacerbation of acute episodes has been identified in bipolar disorder. Prolonged psychosocial stressors during childhood, such as neglect or abuse, are associated with HPA axis dysfunction (Heim et al., 2001). In future years such dysregulation may predispose an individual to affective disturbance, and those who develop bipolar disorder may experience an earlier age at onset, increased rates of self-harm and psychotic symptoms (Post et al., 2001). Likewise acutely stressful life situations and hostility or criticism in a family may trigger episodes in those with an established illness (Johnson and Fingerhut, 2006). In turn, illness in itself is stressful which may lead to further destabilisation, creating the possibility of a self-perpetuating cycle. The degree of negative emotionality expressed by close family members (termed expressed emotion, or EE) has been shown to predict future depressive episodes in patients with bipolar disorder (Yan et al., 2004) and levels of depressive and manic symptoms (Kim and

Miklowitz, 2004, Miklowitz et al., 2005). The high prevalence of bipolar disorder in ethnic minority groups, as demonstrated in recent studies in the UK (Lloyd et al., 2005), may relate to the psychosocial stressors of social isolation and lack of social support often experienced by these groups (Bentall, 2004, Leff, 2001).

Traumatic experiences in childhood have been associated with an adverse course of bipolar disorder and the development of comorbid post traumatic stress disorder (PTSD) in adult life (Goldberg and Garno, 2005). Retrospective studies have shown an association between a history of childhood abuse and an earlier age at illness onset, increased comorbid substance misuse disorders, increased axis I and II comorbidities, and a rapid cycling course (Garno et al., 2005, Leverich et al., 2002).

Theories of the psychology of bipolar disorder have identified factors such as self-esteem and explanatory style that may contribute to mood symptoms. The manic defence hypothesis explains the appearance of manic symptoms as an attempt to avoid the negative and ego-destroying thought patterns associated with depression and anxiety. The ascent into feelings of omnipotence and triumph are thought to overcompensate for feelings of worthlessness and underlying depression which are seen as the backdrop to the manic syndrome (Bentall, 2004). This formulation suggests there is a degree of fragility to the manic state and evidence of negative self-concept or thinking styles should be evident in both patients with mania and remitted patients. There is evidence that patients with bipolar disorder have a negative self-concept, highly variable self-esteem and increased drive even during the remitted state (Bentall et al., 2005, Lyon et al., 1999, Winters and Neale, 1985). Studies using implicit or disguised measures of explanatory style have found that remitted patients tend to attribute negative outcomes to themselves, but positive outcomes to others – a thinking style typical of patients with depression (Lyon et al., 1999, Winters and Neale, 1985). However, this may be better understood as chronic low-grade depression due either

to the debilitating aspects of the illness or due to the physiological processes outlined above rather than as the underlying fuel for mania. Nonetheless, psychological theories of bipolar disorder may help observers understand some of the ideas and beliefs held by those suffering from mania.

COMORBIDITY

Comorbidity is the norm rather than the exception in bipolar disorder. A study of 288 patients with bipolar disorder found 65% had suffered from at least one other (axis I) disorder at some point in their lifetime and one third had at least one current comorbid (axis I) diagnosis (McElroy et al., 2001). The most common comorbid axis I disorders are anxiety and substance misuse disorders, both of which occur in approximately 30-50% of patients with bipolar disorder. In terms of comorbid personality disorders, one study reported a comorbidity rate of 38% in euthymic patients with bipolar disorder (Kay et al., 1999). Those who have comorbidities tend to have had an earlier age at onset and are more likely to experience cycle acceleration and suffer a more severe illness than those without (McElroy et al., 2001).

COURSE AND PROGNOSIS

For most patients bipolar disorder is chronic and recurrent. There is a large amount of variation between individuals in the number of episodes experienced, but the average is 10 (Mackin and Young, 2005). Episodes of mania and depression tend to cluster together so typically patients may experience a number of illness episodes together followed by a more quiescent period and then another cluster of episodes. This pattern with hypomanic and depressive episodes is especially common in bipolar II disorder. The risk of recurrence in the 12 months after a mood episode is especially high (50 per cent in one year, 75 per cent at four years, and afterwards ten per cent per year) compared to other psychiatric disorders.

Furthermore, compared to unipolar depression, bipolar disorder is much more changeable in severity of the mood episode. In those with a recurrent illness pattern, the length of euthymia between episodes may shorten over time suggesting increased frequency of episodes (Kessing et al., 2004). The length of episodes remains fairly constant for an individual over time, although later episodes may begin more abruptly.

NEUROPSYCHOLOGICAL FUNCTION

Over the past twenty years a growing body of literature has identified that many patients with bipolar disorder have significant neuropsychological impairments. These impairments are most pronounced when patients are symptomatic (either depressed or manic). However, deficits have also been reported in euthymic patients. Meta-analyses have indicated the domains showing greatest impairment are declarative memory and aspects of executive function such as control of working memory, mental manipulation, fluency, and set-shifting (Arts et al., 2008, Robinson et al., 2006).

There is no obvious reason to expect neuropsychological impairments in otherwise well patients and they present several conundrums for researchers. One is to understand their functional relevance and whether in the everyday lives of patients these are 'silent' deficits or whether they are a barrier or impediment to successful functioning in work or family roles. Deficits evident in the neuropsychologist's office do not necessarily translate into everyday impairments, as the real world often offers alternative approaches or strategies for overcoming problems or processing information that may not be available during a structured and time-limited formal testing session. Certainly, there is both evidence of a subjective awareness of cognitive difficulties in some patients (suggesting at the very least they are of sufficient degree to be noticeable) (Burdick et al., 2005, Martinez-Aran et al., 2005) and of functional impairment in euthymic patients with bipolar disorder (MacQueen et al., 2001). However, the relationship between subjective and objective

cognitive impairments is weak (Burdick et al., 2005). A recent study in neurological patients indicated that subjective complaints of cognitive difficulties were more strongly related to mood than to level of objective cognitive difficulties (Marino et al., 2009). The same may be true in bipolar disorder and may be related to residual depressive symptoms rather than measurable levels of cognitive impairment. Ascertaining whether there is a relationship between objective cognitive impairment and level of social functioning is necessary to understand the significance of these deficits and to lay the foundation for understanding how their impact can be lessened.

The etiology of the deficits is a further conundrum in the neuropsychology of bipolar disorder. There has been significant debate as to whether neuropsychological dysfunction in bipolar disorder displays the sort of properties that would make it useful for identifying more phenotypically homogenous subgroups of patients for genetics studies. There is a considerable lack of longitudinal studies in this area, therefore it is difficult to know with confidence whether the deficits tend to predate illness onset, co-incide with illness onset, or postdate illness onset. Cross-sectional data suggests that the deficits are present early in the illness and that the verbal memory impairment correlates negatively with number of episodes (especially manic episodes), whereas the executive function impairment is not strongly related to illness features (Robinson and Ferrier, 2006). Meta-analysis of a small number of studies in unaffected first degree relatives of patients with bipolar disorder has indicated that a more subtle degree of impairment is evident in more circumscribed neuropsychological domains than those seen in patients (Bora et al., 2009). This suggests that the scale of the impairment is related to the degree of genetic risk for the condition, and also that there may be a core dysfunction that is unrelated to the experience of symptoms (or treatment). High quality longitudinal studies are essential to address these key etiological issues.

A further conundrum raised by reports of neuropsychological impairment in bipolar disorder is what is core nature of the deficit? As mentioned above, several domains show impaired performance and what remains unclear is whether the processes affected each involve a single impaired process or whether there are several distinct impairments. For example, two of the largest areas of impairment – verbal declarative memory and executive function – may be related. All neuropsychological tests involve multiple cognitive processes, and whilst verbal declarative memory tests draw heavily on memory processes, performance on the supraspan word-list tests used can also be supported by executive resources. One of the aims of the present study is to investigate this further and the issues are discussed in greater depth in Chapter 2 on page 30.

CHAPTER 2: INTRODUCTION TO NEUROPSYCHOLOGICAL FUNCTION IN BIPOLAR DISORDER

The nosological roots of affective disorders were planted in the notion that they were functional pathologies – that is the mood and behavioural disturbances occurred without obvious structural organic cause. Without known brain dysfunction, there was therefore little reason to expect patients to exhibit deficits on neuropsychological evaluation. However, almost fifty years ago Kiloh (1961) described a series of patients with psychiatric illnesses (predominantly depression) whose symptoms mimicked the cognitive changes seen in patients with dementia (Kiloh, 1962). As the cognitive symptoms tended to remit when the depression was treated, Kiloh termed it pseudo-dementia. By the eighties and early nineties methodologically rigorous systematic studies of neuropsychological function emerged contrasting patients with major depression or bipolar disorder with healthy comparison subjects. Although rarely performing as poorly as patients with dementia (Christensen et al., 1997), depressed patients and patients with bipolar disorder showed deficits when acutely unwell with either depression or mania (Clark et al., 2001, Kurtz and Gerraty, 2009, Martinez-Aran et al., 2004b, Zakzanis et al., 1998).

Focusing specifically on bipolar disorder, there is also evidence that cognitive impairment during an acute episode fails to remit fully when symptoms improve. There are in excess of 50 studies that have investigated at least one aspect of cognitive function in euthymic patients with bipolar disorder. Five independent research groups have published meta-analyses which have shown very similar results – euthymic patients with bipolar disorder show deficits in several areas, but most prominently in verbal memory and aspects

of executive function (Arts et al., 2008, Bora et al., 2009, Kurtz and Gerraty, 2009, Robinson et al., 2006, Torres et al., 2007). Visual memory has been less commonly investigated, but in the most recent meta-analysis showed an effect size similar to that for verbal memory. It is worth considering these findings a little further.

COGNITIVE FUNCTION IN EUTHYMIA

The following section is based on a systematic literature search (see Appendix 1 for full details). However, owing to the number of good quality reviews and meta-analyses already available in this area (e.g. Arts et al., 2008, Bearden et al., 2001, Bora et al., 2009, Kurtz and Gerraty, 2009, Robinson et al., 2006, Torres et al., 2007) it was considered more appropriate to provide a selective review of the data most pertinent to the present study.

Table 1.1 below summarises the results of the meta-analyses of cognitive function in euthymic patients with bipolar disorder. The meta-analyses all draw on a common pool of studies and overall there is broad agreement between them. The difference between the highest and lowest reported effect sizes for any given measure vary between 0.01-0.48. In general there is better agreement for the verbal and visual memory measures than the executive and attention measures. On closer inspection of the data, the differences in the attention measures most likely relate to methodological differences in how data from different measures was combined (particularly the continuous performance test measures, which were derived from several different but similar tasks).

The effect sizes are largest for measures of executive function, verbal list-learning and visual memory. The lowest effect sizes are reported for visual copy, immediate verbal memory, and recognition memory. The only test index reported by each meta-analysis to show a large effect size is total learning on the verbal memory list-learning tasks.

Within the executive functions, verbal fluency by letter and categories achieved on

Table 1.1: Comparison of effect sizes (Cohen's d) from the five meta-analyses

Test Index	Robinson 2006	Torres 2007	Arts 2008	Bora 2009	Kurtz 2009
Executive					
Category fluency	1.09	-	0.75	-	0.75
Stroop	0.63	0.71	0.73	0.76	0.75
Trail Making Test B	0.78	0.55	0.73	0.86	0.73
Reverse Digit Span	0.98	0.54	1.02	0.75	0.65
WCST $_{\text{Perseverative Errors}}$	0.76	-	0.72	0.70	0.61
WCST $_{\text{Categories Achieved}}$	0.62	0.69	0.49	0.66	0.54
Letter Fluency (FAS)	0.34	0.47	0.47	0.60	0.51
Verbal Memory					
A/CVLT total learning	0.90	0.81	-	0.85	0.81
A/CVLT short delay recall	0.73	0.74	0.82	0.73	-
A/CVLT long delay recall	0.71	0.72	0.85	0.77	0.78
A/CVLT recognition	-	0.43	-	0.44	-
Forward Digit Span	0.47	-	0.37	0.37	0.41
WMS verbal memory immediate recall	-	-	-	-	0.63
WMS verbal memory delayed recall	-	-	-	-	0.92
Visuospatial					
ROCF copy	-	-	0.22	0.23	0.26
ROCF immediate recall	-	-	0.62	0.59	0.73
ROCF delayed recall	-	-	-	-	0.80
WAIS Block Design	-	-	-	-	0.55
Attention & Psychomotor					
CPT reaction time	0.60	0.62	-	-	-
CPT sensitivity	0.48	0.74	0.58	0.83	0.69
WAIS DSST	0.59	0.79	0.84	0.75	0.66
Trail Making Test A	0.52	0.60	0.58	0.69	0.65

Key: White Small effect size (0.2≤d<0.5) Medium effect size (0.5≤d<0.8) Large effect size (d≥0.8)

WCST, Wisconsin Card Sorting Test; A/CVLT, Rey/California Auditory Verbal Learning Test; WMS, Wechsler Memory Scale; ROCF, Rey-Osterrieth Complex Figure; WAIS, Wechsler Adult Intelligence Scale; CPT, Continuous Performance Test; DSST Digit Symbol Substitution Test

the Wisconsin Card Sorting Test consistently show lower effect sizes (generally d<0.65) than the other executive functions. The remaining measures show a very similar degree of impairment (generally 0.7<d<0.8).

The evidence indicates relatively broad impairment in patients with bipolar disorder, with the most pronounced impairment in verbal memory, visual memory, and several aspects of executive function (e.g. category fluency, inhibition, set-shifting and mental manipulation). Discounting a few extreme results, the largest deficits tend to be in the range 0.7<d<0.85.

COGNITIVE DYSFUNCTION AS AN ENDOPHENOTYPE

There has been much debate about whether cognitive dysfunction in bipolar disorder could be an endophenotype. The observed behaviours and symptoms that form the phenotype of bipolar disorder show significant heterogeneity both between individuals and within an individual over time. Some of this heterogeneity is inherent in the criterion-based diagnostic system used to define the limits of what comprises the illness. Two different individuals with ostensibly the same diagnosis can have remarkably different presentations, to the extent it may be feasible to question whether they have the same underlying illness. The difficulty for research, particularly that which is focused on the underlying genetics, is identifying aspects or traits of the illness that lie somewhere between the genotype and the phenotype – a so-called endophenotype – that can be used to identify subgroups of individuals more likely to share commonalities in their illness pattern or aetiology.

To satisfy the criteria expected of an endophenotype, cognitive dysfunction needs to be demonstrated to be associated with the illness, independent of mood state, heritable, and present in relatives of those with the disorder more commonly than observed in the background population (Hasler et al., 2006). It performs variably on these four criteria.

Firstly, with respect to association with the illness, although there is a large body of evidence (described above) showing that patients with bipolar disorder perform more poorly than healthy controls on neuropsychological assessment, there is no convincing evidence that the deficits observed are specific to bipolar disorder. Patients with schizophrenia (Barch, 2009), obsessive compulsive disorder (Aigner et al., 2007, Segalas et al., 2008), major depression (Austin et al., 1999, Paradiso et al., 1997), or personality disorders (Dinn et al., 2004, Monarch et al., 2004, Ruocco, 2005) also show impairments on many of the same cognitive measures. This may reflect the broad-based sensitivity of many of the neuropsychological measures commonly used, or that similar cognitive processes are affected in a variety of psychiatric pathologies.

Secondly, the deficits are independent of mood state, in that there are a core set of impairments noted in euthymic patients, but symptomatic patients tend to show a more extensive range of deficits than those who are well. There is also some debate about the level of subsyndromal symptoms that are relevant in terms of their impact on neuropsychological function. Studies of euthymic patients rarely recruit individuals with no symptoms at all, but rather those with very low levels of symptoms that – in the context of their illness – reflects a state of relative wellness. However, most studies report significant differences between patients and controls on their symptom levels, even in patients whose scores are nonetheless very low. Although statistical control for symptom levels rarely renders all neuropsychological differences statistically non-significant, a question remains whether statistical techniques can reasonably correct for something that is fundamentally different between the groups rather than a difference that has simply arisen by the misfortune of chance – the former representing circumstances under which covariance techniques are inappropriate (Strauss and Allred, 1987). Analysis of covariance (ANCOVA) was originally designed to increase power to detect between-group differences by reducing within-group variance that had arisen due to chance differences between groups. ANCOVA

therefore assumes random allocation to groups. Patients and controls are not allocated randomly to groups and measures such as symptom scores do not vary randomly between the groups. Instead they tend to vary systematically and are in fact intrinsically related to group membership. Using symptom scores as covariates in these circumstances is generally inappropriate and may under-correct for their 'true' effects, that is statistically significant differences may remain evident when they ought not to.

Thirdly, with regard to heritability, although there is evidence that cognitive function in general shows a degree of heritability, there is only a single study investigating heritability of cognitive function specifically in bipolar disorder (Antila et al., 2007). This study reported a statistically significant degree of heritability for psychomotor speed, working memory, and executive function, but no significant heritability for verbal learning. This is an interesting finding given that verbal learning is one of the areas of greatest impairment in bipolar patients, yet it does not show evidence of heritability. In a similar vein, Szoke et al reported significant familial resemblance between bipolar patients and their first degree relatives for executive function and psychomotor speed measures (no other domains were assessed in this study) (Szoke et al., 2006). Taken together these results provide evidence that cognitive function in bipolar disorder is heritable and some of the areas showing heritability are those where impairment has been found in patients. However, the studies to date have only included patient samples where bipolar disorder runs in the family and have comprised probands and their first degree relatives (with only a small subsample of second degree relatives in the study by Antila et al). Study samples with a greater range of relatedness which also include cases with no prior family history of the illness are necessary to build on these initial findings. Also, one of the most robust areas of cognitive impairment in bipolar disorder – verbal memory – did not show significant heritability. If confirmed in further studies, this anomaly would require investigation and may hint at the possibility that there is more than one process leading to cognitive impairment in bipolar disorder.

Finally, there is evidence that unaffected first degree relatives of patients with bipolar disorder show subtle cognitive impairments when contrasted with healthy individuals with a benign family psychiatric history (Ferrier et al., 2004, Keri et al., 2001, Sobczak et al., 2002, Zalla et al., 2004). The deficits are smaller in magnitude than those found in patient samples and are generally restricted to verbal memory and executive function. However, some studies have investigated individuals below the mean age of onset of bipolar disorder and in any study including high risk groups some of the participants may yet go on to develop the condition. Without a substantial period of follow-up it is therefore not possible to ascertain whether in some of these individuals, the cognitive deficits are a prodromal sign of illness.

All in all, on present evidence alone, there is not strong support for cognitive dysfunction as an endophenotype in bipolar disorder. However, the state of the evidence itself is somewhat lacking. There is a growing trend for studies to divide patient samples on potentially relevant characteristics, such as history of psychosis, substance misuse, level of social functioning and other severity of illness indicators, which is beginning to provide more detail about factors relevant for cognitive functioning in bipolar disorder (Ferrier et al., 1999, Glahn et al., 2007, Martinez-Aran et al., 2008, Van Gorp et al., 1998). Some studies are beginning to delve into greater depth to understand the core processes behind the reported deficits, using novel tasks or novel analysis methods to attempt to tease apart the multiple different processes which contribute to any single neuropsychological task (Glahn et al., 2006, Thompson et al., 2007, Thompson et al., 2006). Although in its infancy, this line of inquiry may ultimately result in the development of more specific test batteries that are better able to indicate the major factors driving impaired performance.

AETIOLOGICAL ISSUES

One of the major issues that remains largely unknown is when cognitive function develops in bipolar disorder. Does it pre-date illness and therefore potentially hold insights into the underlying pathology, or does it develop after mood symptoms and shed light on the consequences of the disorder or its treatment? The gold standard evidence – prospective longitudinal data including the pre-illness period – is rare and so inferences have to be made from a variety of study designs.

LONGITUDINAL STUDIES

PROSPECTIVE STUDY

There has been one longitudinal prospective study of individuals at high risk for mood disorder, which reported that two thirds of individuals who met criteria for bipolar disorder by early adulthood had shown impaired performance on the Wisconsin Card Sorting Test (WCST) when assessed in adolescence (Meyer et al., 2004). In contrast, only one fifth of those who developed unipolar major depression showed impaired WCST performance, which was comparable to the rate found in those who did not develop any major mood disorder. However, this study involved only a small number of participants who ultimately developed bipolar disorder (n=9), two of whom already had bipolar disorder in adolescence. A larger study is necessary to confirm this finding.

RETROSPECTIVE STUDIES

Three retrospective studies conducted in countries with comprehensive population registers used cohorts of conscripts to examine the association between cognitive performance at conscription and subsequent development of psychiatric disorder (Reichenberg et al., 2002, Tiihonen et al., 2005, Zammit et al., 2004). Zammit et al (2004) reported no relationship between IQ and bipolar disorder (whereas low IQ was associated

with later development of schizophrenia). Similarly, Reichenberg et al (2002) reported no relationship between cognitive performance and later diagnosis of nonpsychotic bipolar disorder in a cohort of Israeli conscripts. In contrast, Tiihonen et al (2005) reported that poor performance on the visuospatial reasoning subtest of the Finnish Defence Forces Basic Ability Test was associated with later hospitalisation for bipolar disorder and schizophrenia. However, bipolar disorder was associated with spared or superior performance on the mathematics subtest, whereas schizophrenia was associated with poor performance on this subtest. Performance on the verbal reasoning subtest did not (independently) predict subsequent psychiatric illness. Neither the Zammit et al (2004) study nor the Reichenberg et al (2002) study examined performance subscales of the IQ tests. The relationship reported in the Tiihonen et al (2005) study was specifically with hospitalization for bipolar disorder rather than simply diagnosis of bipolar disorder, which diverges from the other two studies. This may account for the apparent differences between them; those who go on to have episodes severe enough to require hospitalization may be more likely to show cognitive difficulties.

These retrospective case-register studies suffer several limitations - notably the tests used are usually generic tests suitable for mass-administration or for administration with relatively little training; the tests are aimed to identify strengths and weaknesses for channeling recruits into their most appropriate role, they are not designed to detect cognitive impairment and their psychometric properties for that purpose remain largely unexplored; psychiatric symptoms are not thoroughly assessed at the time of testing (beyond medical checks to establish sufficient health for serving in the armed forces) and some individuals may already be showing symptoms or prodromal signs of illness that are not acute or severe enough to be detected. Nonetheless, the naturalistic nature of the data and the large sample sizes make these studies a rich source of information, and have indicated that in generic assessment any prodrome to bipolar disorder is at best subtle and

for the majority of individuals does not represent markedly anomalous cognitive function before the onset of symptoms.

POST-ONSET LONGITUDINAL STUDIES

After onset of illness, is there any evidence that cognitive performance declines over time? There are a small number of test-retest studies that begin to address whether cognitive function shows deterioration over the course of the illness (Balanza-Martinez et al., 2005, Gildengers et al., 2009, Mur et al., 2008). The longest of these - conducted over a three-year interval - reported evidence of cognitive impairment at both baseline and follow-up, but no evidence of significant deterioration between assessments (Balanza-Martinez et al., 2005). In a similar study conducted over two years, again patients showed evidence of impairment but the authors reported no significant deterioration between assessments (Mur et al., 2008). However, there was a significant interaction between group and time for verbal learning, such that patients with bipolar disorder worsened slightly over time whereas controls showed a slight improvement. Given that neither change was significant in itself, the authors dismissed the finding and focused on the persistent and stable executive function impairment seen in the patients. However, the differential trajectory of memory function between patients and controls may be of importance. To clarify whether this was indicative of a difference between the groups, further analysis comparing performance at the endpoint controlling for baseline performance, or analysis of predicted scores at endpoint accounting for baseline performance and practice effects would have been informative. This relatively subtle suggestion of differential memory function in bipolar patients may be indicative of a genuine effect that would be more evident in a larger sample followed-up for a longer period of time.

A few studies have focused specifically on older patients with bipolar disorder. Depp et al (2008) reported no difference in the trajectory of cognitive change in middle-aged and

older bipolar patients compared to healthy controls over 1-3 years of follow-up, but they did note greater within-subject variation in the participants with bipolar disorder than either healthy controls or a comparison group with Schizophrenia (Depp et al., 2008). However, participants in this study were not necessarily euthymic or in the same mood state at the two different testing points (in contrast to some other similar studies). The baseline demographic data indicate that the bipolar patients were experiencing a higher mean level of depressive symptoms with a larger variance than the other groups which may account for the greater variation in performance. In a sample of older patients with bipolar disorder (all over 50 years of age), Gildengers et al (2009) showed evidence of more rapid cognitive decline over the period of three years compared to healthy controls. Using a generic dementia-screening battery administered annually, bipolar patients showed impairment at all time points, and whereas controls showed only a slight decline in performance over the three years, the patient sample showed gradual deterioration year on year. By the end of the study eight patients scored below the cut-off for dementia, in contrast to none of the participants in the control group. This study suffered large attrition, with less than fifty percent of the initial sample remaining in the study by the finial assessment point, which may have influenced the results especially as attrition was higher amongst the relatively younger participants. Likewise, although the average age of illness onset for the whole sample was consistent with that usually seen for the disorder (late-twenties/early thirties), the range indicated there was at least one individual with very late onset (>70 years). Late onset bipolar disorder is associated with greater neurological problems (Fujikawa et al., 1995, Shulman, 1997) and even a small number of participants with late onset could introduce notable bias.

The data so far indicate that over relatively short periods of time there is not a significant change in cognitive function in individuals with bipolar disorder. It remains to be seen whether change is evident over longer follow-up periods. Given the extent of the

impairment shown by patients (between 0.4d-0.8d), if it were assumed that patients start at a comparable level to healthy comparison subjects and show gradual deterioration after the onset of symptoms, then the deterioration within one to three years is likely to be of such marginal size in terms of effect size that it is unlikely to be detectable in the studies conducted to date. To detect such a small change at statistically significant levels would require an unfeasibly large sample size.

CROSS-SECTIONAL STUDIES

Several cross-sectional studies have used correlation or contrasted patient groups with different clinical characteristics to understand how cognitive function relates to the experience of illness.

SINGLE VERSUS MULTIPLE EPISODES

One study contrasted cognitive performance in patients who had experienced only a single episode of bipolar disorder with those who had experienced multiple episodes (Nehra et al., 2006). Both groups showed significant impairment compared to controls, but (counter-intuitively) greater impairment was seen in the first episode patients than the multi-episode patients on two executive measures and one psychomotor function measure. There are several reasons why this might have been. Firstly, all of the first-episode patients, by definition, had a manic episode as their index episode. There is some suggestion that individuals whose first episode is manic have a distinct form of bipolar disorder with a different longitudinal course (Forty et al., 2009, Perugi et al., 2000). Secondly, although all patients in both groups were euthymic when tested, the multi-episode patients had on average 1 month longer in euthymia. This difference was not statistically significant, but there may be an important distinction to make between statistical and clinical significance in this circumstance. First episode patients have to adjust to a new diagnosis and the aftermath

of a manic episode, as well as a new medication regime. There may also be an effect of diminishing marginal returns to recovery, whereby each extra week in recovery brings a proportionately smaller gain in 'wellness'. This suggests that early in the process of recovery, a relatively short time period (such as a month) could be associated with a large change in symptoms or general functioning. If the groups were on different points of this curve, it may have had a bearing on the results. However, even considering both of these factors, there remains evidence that cognitive impairment can be noticed very early in the course of bipolar disorder. This study suggests that, if anything, impairment is worst early on in the illness, and then a process of adaptation takes place. Further similar studies are needed to see if this pattern is replicated.

RELATIONSHIP WITH ILLNESS FEATURES

The approach taken most commonly by studies investigating the relationship between cognition and course of illness is to correlate performance on neuropsychological tests with history of illness variables. In a review of these studies, Robinson & Ferrier (2006) identified significant negative relationships between various features of illness and different aspects of cognitive function. The most consistently reported relationship was a negative correlation between the number of manic episodes and delayed verbal recall. Longer illness duration and more hospitalisations were associated with poorer verbal memory function. Depressive episodes were less consistently related to a broader range of impairments. This may reflect greater heterogeneity in the depression phenotype, or the more consistent relationships with manic episodes may reflect the fact that mania may be more accurately documented and remembered partly on account of the fact it is less common than depressive symptoms, it is often more severe, more life-disrupting, and more likely to lead to hospitalisation than depression.

Measures of executive function were less strongly correlated with illness history variables in contrast with measures of memory function. The authors speculated that executive function may represent a more trait-like aspect of bipolar disorder, evident early in the illness but remaining relatively independent of illness course, whereas memory impairment may relate more closely to the progression of illness.

UNAFFECTED FIRST DEGREE RELATIVES

As outlined above, evidence has indicated mild impairments in verbal memory and executive function in first degree relatives of bipolar patients. From an aetiological perspective, this would suggest it may be a trait marker of bipolar illness also detectable in those sharing genetic vulnerability to the disorder.

SUMMARY

The data outlined above suggest that bipolar disorder most likely shows subtle cognitive differences in executive dysfunction before the onset of illness (differences which may be more evident in those known to be at elevated risk of developing psychiatric disorder) followed by a marked deterioration after the first mood episode to a level which remains fairly consistent, but which shows marginal deterioration – particularly in memory function – with repeated episodes. There is some limited evidence that the rate of cognitive decline accelerates in older bipolar patients.

KEY QUESTIONS

RELATIONSHIP BETWEEN EXECUTIVE AND MEMORY DYSFUNCTION

It has yet to be established whether deficits in verbal learning and executive function are two discrete areas of impairment, or whether deficits in executive function introduce inefficiency in encoding or retrieval processes that impedes effective performance

on verbal memory tasks. It is difficult to test specific cognitive functions in isolation and poor performance on a memory test does not necessary implicate memory dysfunction without a thorough assessment of other abilities. In a sample of patients referred for neuropsychological assessment, Duff et al (2005) reported that executive and memory function showed a strong relationship, sharing 55-60% of variance in a principal-components-type analysis (Duff et al., 2005).

The assessment of memory function in bipolar disorder has most commonly been conducted using list-learning tests. In a study by Tremont et al (2000), a group of participants classified as having significant executive dysfunction were significantly impaired on a list-learning memory task compared to participants classified with minimal executive dysfunction, whereas the groups were not significantly different on a passage-recall test (Tremont et al., 2000). This study suggests that list-learning tasks draw more heavily on executive functions than passage recall. The most recent meta-analysis in patients with bipolar disorder was the first to include passage-recall measures. The results indicate that both immediate and delayed passage-recall are impaired in bipolar disorder, however immediate passage-recall shows a smaller effect size (d=0.63) than immediate list-recall (d=0.81). The pattern of effects is reversed for delayed recall with passage-recall showing a larger effect size (d=0.92) than list-recall (d=0.78).

The extent to which executive strategies are used to support memory performance in list-learning paradigms may be reflected in the degree of ordering of the list shown by participants. This degree of ordering, or subjective organisation (SO), has not been investigated in patients with bipolar disorder. Investigating SO may clarify whether the observed memory impairment is due to underlying executive dysfunction. This is discussed in more detail in Chapter 6 on page 131.

IS COGNITIVE IMPAIRMENT IN BIPOLAR DISORDER MODIFIABLE?

A question of significant importance is whether any of the cognitive impairments observed in bipolar disorder can be reduced. Improved cognitive function has been seized upon as an important outcome of drug studies and there are increasing calls for it to become a focus for intervention. To date, the cause of cognitive impairment in bipolar disorder is unknown and investigations exploring whether or how it can be changed could be highly informative in this regard. A deficit that largely remains despite a variety of techniques designed to ameliorate it has different implications to one that shows a modest to large degree of modifiability. Although this approach alone is unlikely to identify the cause of impairment, the information it provides can be added to information from other sources (e.g. structural and functional imaging, neurophysiological studies) to clarify the picture.

In patients with Schizophrenia several non-pharmacological approaches have been used to attempt to modify cognitive performance. Incentives (Hellman et al., 1998), extended instruction (Hellman et al., 1998), and combinations of interventions amalgamated into an extended cognitive remediation therapy programme (van der Gaag et al., 2002, Wykes et al., 2007) have all been tried, often with limited effect. So far there has only been one study of a non-pharmacological intervention designed to modify cognitive function in patients with bipolar disorder (Deckersbach et al., 2009). This is discussed in more depth in Chapter 7 on page 177.

RELATIONSHIP WITH SOCIAL FUNCTIONING

The above review has identified a number of areas of neuropsychological dysfunction in patients with bipolar disorder. It remains unknown what significance these impairments have to patients in their everyday lives in terms of their ability to function. In

patients with Schizophrenia, certain aspects of cognitive function are closely related to psychosocial functioning (Bowie et al., 2008, Green, 1996).

In bipolar disorder, the relationship between functioning and cognition has been explored in a very limited manner. Psychosocial functioning is a multi-faceted construct which is difficult to capture adequately with the use of a single measure, and most studies conducted so far have simply correlated scores on a global functional outcome measure with scores on neuropsychological tests. There is evidence of a relationship in a very general sense (Dickerson et al., 2004, Martinez-Aran et al., 2004a, Martinez-Aran et al., 2007), but no understanding of which aspects of functioning relate most closely to which aspects of cognition. If remediating cognitive deficits in patients with bipolar disorder is to become a realistic focus for intervention, then it is essential to develop a greater understanding of the degree and type of functional improvement that could be expected. This would clarify the value any successful intervention would be likely to have for the individual involved and also identify which cognitive areas should become the major targets. These issues are discussed further in Chapter 4 on page 76.

CHAPTER 3: INTRODUCTION TO EMOTION PROCESSING IN BIPOLAR DISORDER

INTRODUCTION

The central role of mood and affect in establishing a diagnosis of bipolar disorder has not been reflected in the balance of research into the illness. Despite mood change being a key feature of the disorder, there has been a relative lack of research into the impact that mood change has on the processing of emotional (or other) information and on social functioning. Understanding more about whether people with bipolar disorder show differences in how they interpret emotional stimuli or how the processing of emotional stimuli influences other cognitive processes may provide important clues as to the underlying pathology in bipolar disorder. Furthermore, given the central role emotions play in navigating the social landscape, deficits could show a relationship with social functioning, as is the case for individuals with schizophrenia (Kee et al., 2003).

Before reviewing the published studies of emotion processing in bipolar disorder, it is first necessary to consider what is meant by mood and emotion, and to outline briefly different psychological models of emotion that guide the approach to and interpretation of research in this area.

MOOD

Exactly what constitutes a mood state is difficult to define. However, at a basic level, mood states alter an individual's sensitivity to a range of different stimuli and facilitate the experience of some emotional states over and above others (Evans, 2001). The effects of

mood states are not limited to changes in subjective experience or feelings. Their effects are manifold, including (but not limited to) effects on cognitive processing, physiology, energy balance, and motivation.

One of the defining features of moods that is broadly agreed upon is that they are diffuse or global states, and, related to this, they therefore lack intentionality (Siemer, 2009). Moods are not directed at a specific object, instead they are general, free-floating states that may have no obvious trigger or cause. Moods can arise independently of a specific triggering event or appraisal or remain after an emotion-eliciting stimulus has ended. There are opposing views about the nature of moods. As described by Siemer (2009), 'mood as core affect' theorists conceptualise moods as the feeling component of emotions and hypothesize that moods are essentially the same as emotions, but lack some of the elements of an emotion (e.g. a cognitive appraisal). One opposing school, the 'mood as a temporary disposition' theorists, propose that moods reflect a tendency to respond to a stimulus in a particular way, a way that is consistent with that mood (Siemer, 2009). Moods and emotions are conceptualised as forming a continuum rather than two qualitatively different states. The value in this latter approach is that it provides a link between mood and emotion. In the present context, it is also a valuable rationale for investigating mood-based processing biases in patients with mood disorders. 'Mood as a temporary disposition' theorists describe moods as "temporarily heightened dispositions to have or to generate particular kinds of cognitions, specifically to make particular kinds of emotion-relevant appraisals" (Siemer, 2009, p.257) suggesting cognitive biases are to be expected in different mood states.

The 'mood as a temporary disposition' theory is also consistent with evolutionary approaches to understanding emotion. From an evolutionary standpoint, it has been hypothesized that the function of mood states is to facilitate a response to environmental stimuli in such a way as to minimise the discrepancy between the expected and actual

payoff of any given action (or inaction) (Tooby & Cosmides, 1990). In other words, signals from the environment (where 'environment' includes both factors internal and external to the individual) provide information about which lines of action are likely to lead to valued outcomes. As all actions are not equally likely at any given point in time, by this view a mood state serves to increase the probability of producing actions more likely to lead to desired outcomes and to inhibit actions likely to prove costly. This maximises the chances of reaping the benefits of the current circumstances. For example, significant losses (real or imagined) are frequently triggers for sad mood, especially interpersonal losses. The experience of the loss suggests that continued investment or engagement in behaviours or relationships associated with the loss are unlikely to lead to a positive (i.e. survival enhancing) outcome. The environmental cues are indicating that high investment in, for example, interpersonal situations is likely to expend more energy than it is to reap reward. As such, sad mood facilitates withdrawn behaviour and facilitates the experience of a distressed emotional state, which may create sympathy in others that in turn may assist in ameliorating the effects of the initial loss. In this way, the sad mood state has increased the likelihood of the individual engaging in behaviour that is more suited to the prevailing circumstances.

Evolutionary approaches understand behavioural patterns evident in the present as adaptive solutions to constellations of events that occurred repeatedly in the ancestral environment. Over thousands of exposures to the same constellation of events, an adaptive solution evolves owing to the enhanced survival and reproduction of those individuals who navigated the situation most successfully. However, in the modern environment there is no guarantee that a) the evolved 'solution' is still adaptive, or b) it will not malfunction or be inappropriately activated. In the case of mood disorders, the persistence of negative or positive mood states and the associated heightened emotionality (reflected in the experience of distress and anger/irritability during depressed moods, or joy and euphoria

during manic moods) reaches an extent that – on an individual level at least – is no longer functional or adaptive.

EMOTION

If mood states form the backdrop of the emotional landscape, then emotions themselves are the foreground features. Although equally difficult to define, emotions are generally considered to be shorter lasting than mood states (lasting seconds to a few minutes), and tend to be triggered by an (internal or external) environmental event of major significance to the individual (Scherer, 2000). This sense of intentionality – that emotions are 'about something' – is often used as the key differentiator between mood and emotion.

There is significant debate about the neural, physiological, physical, behavioural, and cognitive features that constitute an emotional experience and which differentiate emotions from other internal states such as mood, drive states, attitudes, or character traits. Izard (1993) identifies three basic characteristics of an emotion that generate broad agreement amongst most emotion researchers: 1) Emotions involve particular neural processes, i.e. they are not a general process involving the entire brain. Different theorists have proposed different specific processes, but as of yet no agreement has been reached. 2) Emotions involve an expressive or motor component. Some theorists have proposed specific patterns of facial expressions that are characteristic features of particular emotional states (Ekman and Friesen, 2003). However, it is acknowledged that the expression of emotional states can, under certain circumstances, be voluntarily suppressed (Ekman and Friesen, 2003). Therefore, although the expression is absent, other features of the emotion are still present. These theorists would agree that at the very least, emotions involve efferent activity in the central nervous system which may or may not be translated into actual motion. 3) Emotions register in consciousness. The most straightforward understanding of this is that emotions often involve a subjective feeling state. However, Izard lists several other ways in which

emotions may register in consciousness irrespective of a change in subjective feeling, for example by changing the motivational state, creating action readiness, generating an action tendency, creating perceptual selectivity/bias, or acting as cues for cognitive processes (Izard, 1993). There is considerable debate about whether emotions necessarily involve a conscious component (LeDoux, 1998). Experimental paradigms using subliminal stimuli have demonstrated that emotions can be influenced without conscious awareness (Zajonc, 1980), and indeed some would argue that this is when emotions are most vulnerable to influence (LeDoux, 1998). Certainly evidence from the evolutionary tradition would support the view that emotions do not always require conscious awareness, since conscious awareness (and the verbal reporting of emotional states by humans) is a relatively recent addition in evolutionary development that post-dates the evolution of emotion.

If mood disorders are associated with abnormal persistence of dysfunctional mood states, and if mood states genuinely do alter the susceptibility to experiencing emotions and create a selective or biased way of processing information, it could be expected that individuals who experience pathological mood states will show differences in the way they process environmental emotional cues.

MODELS OF EMOTION

There are a number of different families of psychological models of emotion, each of which approaches the definition and understanding of emotion from a different standpoint. The models vary from one extreme of viewing all emotions as biologically hard-wired and honed by evolution through to defining them as entirely socially-engineered constructs. Scherer (2000) summarises four different psychological approaches to emotion (Scherer, 2000):

1) *Discrete or categorical models* – identify a limited number of discrete emotions, each with characteristic identifiable features (e.g. facial expression). For example, following an extensive body of research into human emotion conducted by Paul Eckman, he proposed that there are six basic emotions each with a distinct recognisable external expression: happiness/joy, sadness/distress, anger, fear, disgust, and surprise. Models in this tradition emphasise the evolutionary origins of emotions, and focus closely on their biological underpinnings and neurocircuitry (Darwin, 1872, Ekman and Friesen, 2003). These theorists focus on the production end of emotion in terms of the external signs of the emotional state (such as facial or vocal expression) that are detectable to an observer. They are criticised for being reductionist and restrictive, limiting the range of emotional experience to a few discrete categories, and for a degree of naivety in assuming that all emotional signals are honest. This assumption has the benefit that research into the decoding of emotional signals (as much of the work in this tradition has involved) can be generalised to tell us something about the production of those signals, but fails to acknowledge that senders and decoders of emotion can often have very different motives (Russell et al., 2003).

2) *Dimensional models* – identify a limited number of factors or dimensions that underlie all emotions. Specific emotions are differentiated by their position on each of these dimensions. The most well-known dimensional models include two dimensions, one capturing valence (positive-negative) and one capturing arousal or activation (Russell, 1980). By this model, emotions such as sadness would be represented by negative valance and low arousal, whereas anger would involve negative valance and high arousal. The benefit of dimensional models is their parsimony (they permit the explanation of a wide variety of emotions with very few underlying dimensions). They are criticised for being difficult to test as they rely heavily on essentially unseen constructs that have been derived through statistical techniques (principally some type of factor analysis). Additionally, the number and labelling of the relevant dimensions is not agreed upon.

3) *Meaning-oriented models* – are founded on the belief that emotions are sociocultural constructs and the language used by different cultures to describe emotional experience reveals information about the underlying structure of emotional processes (Harre, 1986). Tied in with meaning-oriented models is the social constructivist perspective, which states that emotions are determined by society and the biological aspects of emotions are secondary to the meaning that the eliciting stimulus takes on within the sociocultural context in which it occurs. Meaning-oriented models have been criticised for focusing too heavily on the subjective feeling component of emotional experience, and essentially breaking the evolutionary link between human emotions and emotions in pre-linguistic species.

4) *Componential models* – place cognition at the centre stage of emotion elicitation. By these models affect is post-cognitive, occurring only after the relevant stimulus has undergone a degree of cognitive processing. All emotions are hypothesized to be triggered by a cognitive evaluation of the relevant stimulus, which places it in a context with regard to its significance for the individual and its likely impact. The emphasis is on the circumstances surrounding emotion elicitation, with less attention paid to *which* emotion is elicited in any particular situation or how the emotion experienced is 'selected' and then 'produced'. Componential models have received much criticism for weighting the cognitive aspect of emotion so heavily and for introducing a 'black box' of processes between stimulus perception and emotional response that cannot be directly measured - much of the cognitive activity hypothesized to take place merely has to be inferred. Some cognitivists have also been criticised for broadening the definition of 'cognitive' to include pure sensory input and other mental processes generally not classified as 'cognitive' (Izard, 1993, Zajonc, 1984). Two prominent theorists engaged in a very public debate towards the end of the last century about the relationship between cognition and emotion. At one extreme, Lazarus argued a cognitive evaluation is always required before an emotion can be experienced (e.g.

Lazarus, 1984). On the other hand, Zajonc argued that – at least in some circumstances – cognition and emotion can be completely independent, with some emotions occurring in the absence of cognitive processing (e.g. Zajonc, 1984). By and large this debate has not been resolved, although the ensuing decades have clarified some of the illusory differences between the two stances created by different usage of the same terminology, and several intermediary positions have evolved (see Izard, 1993). There may be different routes to emotion. Some stimuli may have preferential access via the 'quick' evolutionary route (i.e directly to the amygdala) resulting in emotion elicitation with very little (or no) input from higher-level processing. In other instances, emotions may only be elicited after more conscious and deliberative prefrontal processing has taken place (LeDoux, 1998).

This brief summary of the various models gives an indication of the variety of theoretical perspectives available, each placing emphasis in a different place and with theorists from opposing schools often drawing definitional boundaries in subtly but importantly different places. This makes it difficult to interpret research conducted under one tradition through the eyes of an alternative tradition and serves to highlight differences between models at the expense of identifying their commonalities. As such, there is not yet any convincing evidence to support one model over and above any of the others (or any clear notion what such evidence would look like), but addressing this issue is far beyond the scope of the current project. For the present purpose the precise mechanism underlying the detection and interpretation of emotion is less important than identifying whether there is reason to believe that there is some deficit or dysfunction in this system to begin with. Identifying whether there are differences between patients with bipolar disorder and healthy controls in the way they interpret or respond to emotional signals is a first step towards that end.

EMOTION, COGNITION AND AFFECTIVE DISORDER

The relationship between emotion and cognition is one of significant relevance for the investigation of emotion processing in affective disorder. The view that emotions stand in isolation from other cognitive processes has never been strongly supported, but, as described above, the direction and reciprocity of the relationship continues to be debated.

With respect to mood disorders, the issue is of key importance. As described above, some models of mood place specific emphasis on the fact that mood states dispose an individual to particular types of cognition. Mood states that are pathological in nature, severity or frequency (as are seen in mood disorders) may therefore be associated with significant cognitive biases. Biased cognitive appraisals could in turn perpetuate dysfunctional emotional states creating a positive feedback loop.

In patients with bipolar disorder it is beneficial to explore deficits or biases in the conscious interpretation of emotional displays (e.g. facial or vocal emotion recognition paradigms) as well as the effects that emotional stimuli may have on behavior even when conscious processing of the stimulus is not necessary (as in implicit emotion tasks; see below for examples). The former identifies deficiencies in the decoding of emotional signals that may play an important role in subsequent appraisals. The latter identifies whether emotional stimuli engage processing resources differently in patients, which may subsequently impact on the resources available for other types of processing.

DEFICIT OR BIAS

One question of particular importance is whether emotion processing differences in patients with bipolar disorder represent a deficit or a bias. For example, if patients show worse performance in an emotion recognition paradigm, are the pattern of errors simply random, or do they indicate a consistent misinterpretation that would be consistent with a

biased interpretation of incoming information? In terms of implicit emotion tasks, do patients show greater interference when stimuli of a particular emotional valance are presented, suggesting greater difficulties ignoring these emotions?

Research into major depressive disorder has attempted to identify whether patients show a negative bias in the way they process emotion. The notion of a negative bias in depressed individuals has reassuring face validity – patients with low mood attend selectively to negative information, which worsens mood and so on. This positive feedback loop could explain why negative mood states persist for some individuals. Cognitive models of depression, such as Beck's formulation (Beck, 1979), suggest that dysfunctional negative self-beliefs lie at the core of depression and incoming information is processed in a biased manner in line with underlying core beliefs. Experimental investigation of emotion processing in depressed individuals would therefore be expected to highlight differences in the interpretation of or response to negative information in people who are depressed. Furthermore, individuals who have experienced multiple episodes of depression may continue to show abnormalities even when they are well, indicating ongoing vulnerability to further depressive episodes.

Before reviewing investigations of emotion-processing in bipolar disorder, it is helpful to consider the different types of experimental paradigms that have been used to investigate emotion processing in mood disorders.

EXPLICIT MEASURES

These measures require the participant to make an explicit judgment about the emotional content of the presented stimulus. Although a number of paradigms exist, the most-commonly used in bipolar disorder are facial expression recognition paradigms. The participant is shown a picture or movie of someone showing an emotional expression and the participant is asked to identify which emotion they think the person is expressing. Vocal

expression recognition paradigms, which follow a similar format, have been used less-frequently in patients with bipolar disorder, but are considered in more depth below.

FACIAL EXPRESSION RECOGNITION

No single paradigm has been selected as optimal for assessing facial emotion recognition and many different tasks have been used that vary in the selection of emotions, the stimuli used, the task format (for example some involve labelling the emotions, others judging the intensity of the emotions, others matching emotions). The response options are usually offered in a multiple-choice format, which has been shown to produce very similar results to free-response formats (Rosenberg and Ekman, 1995).

One major difference between paradigms has been the use of dynamic displays of emotion expression versus still facial photographs. Early studies of facial expression recognition used still photographs, but led to concerns over the ecological validity of such stimuli. In everyday life it is common to see an expression formed, rather than see it appear suddenly and completely. Some authors have used imaging software to develop moving images from two still images of an individual starting with a neutral expression and gradually displaying the emotion (e.g. Montagne et al., 2007). The morphing procedure involved mapping almost 180 points of the face on the neutral image and mapping the same points on the endpoint image, then using graphical software to move the points from the neutral start point to the 100% intensity expression in 5% steps. Showing the images in sequence gives a relatively seamless sense of motion from the start to the endpoint. Additionally, emotions can be displayed up to any of 20 different percentages of full intensity (e.g. 50%, 85%, 20%) which allows exploration of which point in expression formation the emotion can be identified.

Although intended to increase ecological validity, there are several issues with this technique. Using 5% steps assumes all parts of the face move at an equal rate when an

expression is formed. This may not be the case. It also assumes all emotions are produced at an equal rate, which may not be true. Additionally, expressions shown at lower intensities are shown on screen for a shorter time than those of a higher intensity (as fewer images are played in the sequence). Repeated displays of the higher intensity expressions may, over sufficient exposures, lead the participant to learn how the expression would have continued to change after it has been stopped. In essence, the participant would then be gauging the expression from a 'forwards prediction' of how they think it would end up, rather than judging it from the observed endpoint. One study comparing dynamic and static displays of facial emotional expressions reported no differences between recognition accuracy for dynamic and static displays (Katsyri and Sams, 2008).

VOCAL EMOTION RECOGNITION

Facial expressions represent only one channel of communicating emotional information and will do so effectively in a limited set of circumstances (e.g. when in relatively close viewing distance from an observer). Vocal expressions, however, can be effective even when the person expressing the emotion cannot be seen. It could be the case, given that facial and vocal expressions potentially serve different functions, that a deficit in emotion recognition may be general to all modalities, or may be present in some but not others.

Research into vocal emotion recognition has received less attention than facial expression recognition in general and also in the field of affect recognition within bipolar disorder. Unique features of the acoustic signal that are relevant for emotion identification have been difficult to identify and many components of the vocal signal seem to be involved (Pittam and Scherer, 1993). Much of the work assessing identification of emotions from vocal information has used a similar paradigm to that used in facial expression identification – playing a sound clip and offering a selection of discrete emotion labels for the participant

to use to identify the emotion. Sound clips are usually derived by asking actors to read statements with a specific emotional intonation. Work using clips derived from naturalistic settings is rare due to the methodological difficulties and lack of control over the stimuli (Pittam and Scherer, 1993). Accuracy (in terms of percentage of presented items correctly classified as the relevant emotion) for vocal recognition is generally lower than that for facial expression recognition, although far exceeds chance levels (Scherer, 2003). Vocal emotion research has generally used stimuli depicting similar basic emotions to those used in facial expression recognition research. The average accuracy level of different vocal emotions varies markedly, with anger generally showing high agreement and disgust being poorly recognised (Pittam and Scherer, 1993). This may reflect that the different basic emotions do not all have a natural vocal expression, or that the role of vocal expression in different emotions varies. For example, the poor recognition of disgust may reflect the fact it is rarely naturalistically expressed over the course of a sentence (the format of the stimuli used in most tasks) and may instead be expressed in short vocal bursts (Pittam and Scherer, 1993). There are also sociocultural and idiosyncratic influences on vocal expression of emotion, which contribute to large variability in the signal produced and may make the production of standardised stimuli especially difficult.

These issues with both facial and vocal emotion recognition paradigms notwithstanding, one further drawback of explicit measures is their vulnerability to demand characteristics. Participants may have preconceived ideas as to how depressed individuals 'ought' to respond which may influence performance.

IMPLICIT MEASURES

In navigating the social world, it is rarely necessary to identify another's emotions in an overt and explicit way, such as is required in explicit emotional identification tasks. However, it is often necessary (or useful) to develop an awareness of another's emotions

and use that in order to plan behaviour. This latter process is likely to vary – both between and within individuals – in the extent to which it occurs under conscious control or awareness. Biases or deficits in this aspect of emotion processing could have noticeable consequences for an individual, for example in how distractible they are when faced with emotional stimuli that are particularly potent for them, or in adapting to and accommodating for others' emotions in social relationships in general.

Measures of implicit emotional processing focus on changes in performance and behaviour that occur as a result of exposure to an emotional stimulus, although the task itself does not require the participant to identify (or necessarily attend to) the emotional content. This usually involves incorporating emotional stimuli into a cognitive task in which explicit identification of the emotion of the item is irrelevant for task performance. Depending on the paradigm, the emotional stimulus may be presented subliminally, or supraliminally. The emotional content of the stimulus is not necessarily masked or hidden in any way from the observer. However, explicit identification of the emotional aspect of the stimulus is not necessary for the task to be performed. By looking at performance on the task in terms of accuracy or reaction time, it is possible to gauge what and how much impact the emotional material had even though conscious processing of the emotion was not necessary.

Several different cognitive paradigms have been adapted for this purpose. Two key areas of interest are whether emotional stimuli cause greater interference in cognitive processing than neutral stimuli and whether emotional stimuli capture or hold attention more so than neutral stimuli. The first of these questions has been addressed using the Stroop test. In the 'emotional' Stroop test, participants are presented with emotional words printed in different colours and, like the standard Stroop, have to read aloud the colour of the word. Participants who have greater difficulty suppressing the emotional content of the

stimulus are hypothesized to show a slower colour-naming time due to response competition/resource constraints. The second of these questions has been addressed using attentional probe tasks. Different variants have been used. In the standard dot probe task two words are presented either side of a central fixation and followed by a single target. The premise underlying the task is that response times to the target are faster if it appears in the same position as the word that had captured the participant's attention. By pairing positive or negative emotional words with neutral words it is theoretically possible to identify biases in attention to words of different valences. In a relatively recent extension of the emotional dot probe paradigm, Koster and colleagues returned to the original Posner-style cuing paradigm (Posner et al., 1980) and presented a single emotional stimulus in one of two possible locations followed by a target dot either in the same location as the cue stimulus, or in the alternative location (Koster et al., 2005). The authors argue that this modification permits examination of both attentional engagement *and* attentional disengagement, whereas the double-cue paradigm is too crude to examine these individual components of attention. Using this paradigm, Koster et al (2005) reported that dysphoric individuals showed impaired disengagement from negative words, potentially indicative of difficulties 'unhooking' attention from negative stimuli (Koster et al., 2005).

PSYCHOLOGY OF BIPOLAR DISORDER AND THE MANIC DEFENCE

Psychological models of bipolar disorder, although in their infancy (Power, 2005), may hold important implications for emotion processing in patients with bipolar disorder.

Following the success of Beck's cognitive model of depression, several theorists have attempted to apply the model to bipolar disorder simply reversing the content of the dysfunctional beliefs from negatively to positively themed, and altering the triggering events from loss-related to goal-attainment-related (e.g. Lam et al., 1999, Leahy, 1999, Newman et al., 2002). This would imply that the negative biases hypothesized to exist in depressed

patients would also occur in depressed bipolar patients, and positive biases would occur in the manic state. However, this model cannot explain which – if any – biases would be expected when patients are euthymic (unlike depression, where negative biases are predicted to remain as vulnerability factors).

Although there are other candidate models that will not be discussed here, one that has particular relevance for emotion-processing findings in bipolar disorder is the so-called 'manic defence hypothesis' , or its less psychodynamically-rooted modern incarnation the 'depression avoidance hypothesis' (Jones and Bentall, 2008). By this model, manic symptoms emerge due to the use of coping mechanisms involving risk-taking and distraction in order to deal with low self-esteem and low mood (Winters and Neale, 1985). As the coping strategies are dysfunctional, they sometimes precipitate a spiral into mania, but at other times are ineffective at preventing the slide into depression. This approach allows room for the observation that manic and depressive symptoms can occur together, and that mania is often a very fragile shroud for the depression and low self-esteem that lie beneath.

In investigating this hypothesis, Lyon et al (1999) identified an inconsistency between explicit measures of self-esteem in manic patients and implicit measures. On explicit measures (those asking patients to rate their own self-esteem), manic patients showed elevated levels of self-esteem. However, on implicit measures (those where esteem is measured indirectly, such as the emotional Stroop test), their performance was more similar to depressed patients (Lyon et al., 1999). Although studies of this nature in acutely ill patients are in their infancy and findings await independent replication, even less is known about the presence of biases during euthymia. Are similar inconsistencies between explicit measures and implicit measures evident in euthymic patients and do the findings generalise to emotional stimuli not related to self-esteem?

EXPLICIT EMOTION RECOGNITION IN BIPOLAR DISORDER

Patients with major depression show differences in explicit emotion tasks consistent with a mood-congruent negative bias (Gur et al., 1992). However, findings on implicit measures have been less supportive of an automatic attentional bias towards all negative stimuli in general. An automatic attentional bias to negative stimuli is reported in anxious individuals (especially for stimulus presentations between 100ms-500ms) (Mathews and MacLeod, 1994), but in depressed individuals a bias is only evident in tasks with longer stimulus presentation times or in which elaborative processing is possible. Additionally, the processing bias is most evident in depressed individuals when the emotional stimuli are self-descriptive negative interpersonal trait terms rather than general negative information, especially if the task has required processing of the terms in relation to the self (Mogg and Bradley, 2005).

FACIAL EXPRESSION RECOGNITION

Most of the explicit emotion-recognition tasks in bipolar disorder have used facial expression recognition paradigms. There have been several studies investigating facial emotion processing in patients with bipolar disorder. A summary of the findings is presented in Table 3.2 below. Results indicate some differences in the explicit processing of facial expressions, however there is no consistent pattern.

MANIC PATIENTS

Of five studies including a sample of manic patients, two reported deficits in emotion labelling (Getz et al., 2003, Lembke and Ketter, 2002). In both studies patients made more errors than controls, with one of the two studies reporting that manic patients had difficulties correctly identifying fear and disgust, mistaking them for surprise and anger respectively (Lembke and Ketter, 2002). The same difficulty was not evident in a sample of

euthymic bipolar patients included as a comparison group. Two of the other studies reported a difference in the manic patients' appraisal of the intensity of expression – patients rated sad faces as less sad than control subjects, but no differences were noted in the appraisal of happy faces (Chen et al., 2006, Lennox et al., 2004). One study reported a non-significant reduction in accuracy in patients with manic symptoms across emotions in general (Gray et al., 2006). This study also included a sensitivity task in which patients had to adjust the intensity of emotion in the face to a point at which they could only just accurately identify the expression. For each of six emotions, patients with manic symptoms were able to identify the expressions at a lower intensity level than controls, indicating greater sensitivity to emotions in general. However, none of the differences was statistically significant on its own, it was only when taken together for all emotions that the findings suggested greater sensitivity in the patient sample. This finding is in contrast with the above studies – the appraisal results described previously suggest that manic patients would require a greater intensity of sadness in the face to identify the emotion. However, patients in the Gray et al study were only experiencing manic symptoms rather than a threshold hypo/manic episode and many were also experiencing depressive symptoms such that the depressive and manic samples of patients overlapped by 6 individuals. The mixed nature of their symptoms may have influenced the results.

DEPRESSED PATIENTS

Two studies included a sample of depressed patients. One reported no differences between patients and controls in the appraisal of the intensity of sad, happy and fearful faces (Chen et al., 2006). The other study reported findings consistent with a mood-congruent bias in emotion processing (Gray et al., 2006). Patients showed a lower level of accuracy in identifying all expressions, although this was only significant when taking all emotions together (no single emotion showed a significant difference between patients and

controls). In the sensitivity paradigm (described above), the patients with depressive symptoms showed reduced sensitivity for happy faces (needing a greater intensity of emotion to identify the expression), but increased sensitivity for all negative emotions taken together.

EUTHYMIC PATIENTS

Five studies included a euthymic or predominantly well sample of patients. One reported no differences in the recognition or discrimination of facial expressions (Vaskinn et al., 2007). One reported no deficits in emotion labelling in euthymic patients with bipolar I disorder, but an enhancement in the recognition of fear by patients with bipolar II disorder (Lembke and Ketter, 2002). One reported enhanced recognition of disgust, which was not accounted for by a general response bias (Harmer et al., 2002). A third study reported deficits in affect matching in bipolar patients (i.e. identifying whether two different faces are showing the same facial expression) (Bozikas et al., 2006). However, no details were provided about whether matching for each of the different emotions showed an equal degree of impairment. Also, this study used images of children rather than adults without explicitly controlling for possible relevant demographic differences between groups (i.e. the number in each group that had at least one child). One further study reported impairment in the recognition of fear, however differential deficit analyses indicated that the recognition of fear was not statistically worse than the recognition of other emotions and therefore the finding should be treated as preliminary (Venn et al., 2004).

CONTROL TASKS

As faces are complex visual stimuli, it is not immediately obvious that differences in facial emotion processing are not due to difficulties processing complex images. However, none of the studies that included control tasks requiring the ability to process faces without

necessarily attending to the emotion (such as identity matching) reported any deficits in bipolar patients (Bozikas et al., 2006, Getz et al., 2003, Venn et al., 2004).

VOCAL EMOTION RECOGNITION

To date there is only one published study of vocal emotion recognition in bipolar disorder (Vaskinn et al., 2007). Using the face/voice emotion identification and discrimination paradigm, Vaskinn et al (2007) reported no deficits in bipolar patients in either identifying emotional intonation or in discriminating between emotions in vocal recordings. The patients in this study were predominantly well (a small proportion of the sample had elevated levels of depression), which may account for the absence of a significant difference. However, the task used did not present an equal number of stimuli for each emotion and only contained a small number of items.

SECONDARY EMOTIONS

Much of the work to date examining explicit emotion recognition has involved stimuli depicting the six basic emotion categories proposed by discrete emotion theorists and derived from evolutionary approaches to emotion (e.g. happy, sad, fear, anger, disgust, and surprise in the nomenclature used by Ekman & Friesen 2003). However, in everyday life the most commonly-observed emotional states involve more complex secondary or 'social' emotions such as guilt, worry, jealousy, love, compassion, confusion etc. These emotions are less about immediate survival and more about successfully navigating the social landscape and sharing a common understanding of how the current circumstances have been interpreted.

Table 3.2: Summary of studies of facial expression recognition in bipolar disorder

Study	Bipolar patients n (age)	% female	Mood state	In-patients (%)	On meds (%)	Control group n (age)	% female	Task(s)	Results
Harmer et al (2002)	**20** (37.8±11.2)	50	Euthymic (HamD<8 & YMRS<8)	0	90	**20** (37.7±17.0)	35	Pictures of Facial Affect	Bipolar patients significantly better at recognising disgust than controls; not due to a general response bias; patients were non-significantly slower at identifying all emotions (except disgust)
Lembke & Ketter (2002)	**8 BDI** (NR±NR)	NR	Manic (YMRS≥20)	100	NR	**10** (NR±NR)	NR	Forced-choice expression recognition task (6 possible response options)	Manic patients were impaired in the recognition of fear & disgust compared to controls and euthymic patients; tended to mistake fear for surprise & disgust for anger; BDII patients were significantly better at identifying fear than the other patient groups – not due to response bias
	8 BDI (NR±NR)	NR	Euthymic (HamD<10 & YMRS<10)	0	NR				
	8 BDII (NR±NR)	NR		0	NR				
Getz et al (2003)	**25 BDI** (25.3±8.4)	52.2	Manic /mixed	100	96	**25** (25.3±7.4)	64	Facial affect matching and facial affect labelling	Patients showed impairment in the facial affect labelling task only
Lennox et al (2004)	**10 BDI** (32.6±10.7)	60	Manic	100	100	**12** (37.3±12.8)	16.7	Shown happy or sad faces at 4 different intensities (0%, 50%, 100%, 150%) and asked to rate how happy/sad the face is	Manic patients rated sad faces as significantly less sad than controls; no differences noted for the ratings of happy faces
Venn et al (2004)	**14 BDI & 3 BDII** (44.4±13.2)	58.8	Euthymic (HamD≤7 & YMRS≤7)	0	100	**17** (43.8±13.9)	58.8	Shown moving images of people producing happy, sad, fearful, angry, disgusted and surprised expressions. 1) Asked to identify the expression (identification); 2) Asked to adjust the image to the point at which they could just identify the emotion (sensitivity)	No differences in sensitivity; patients significantly worse at identifying fear, but this was not a differential deficit
Bozikas et al (2006)	**19 BDI** (39±11)	57.9	Euthymic (MADRS≤8 & YMRS≤8)	0	100?	**30** (38±10)	50	Kinney's Identity Matching Test & Kinney's Affect	Patients significantly impaired on the affect matching test; no impairment in

								Matching Test	identity matching
Chen et al (2006)	**8** (41.9±12.1)	37.5	Depressed	12.5	100	**8** (38.8±12.5)	75	Explicit judgement of expression intensity for sad, fearful, & happy faces (as Lennox et al above); parallel implicit task – same faces shown at 4 different intensities of 3 colours and subjects asked to rate the intensity of the colour	Nonsignificant trend evident for manic patients to underestimate the intensity of sadness; no significant differences in the subjective intensity ratings of the depressed patients
	8 (39.0±13.4)	0	Manic	100	100				
Gray et al (2006)	**14** (45.1±13.82	71.4	Depressed	0	100?	**21** (46.9±13.4)	66.7	Task as per Venn et al (2004) above	Depressed patients were less accurate at emotion identification in general and showed reduced sensitivity to happy faces but increased sensitivity to negative emotions. Manic patients showed non-significantly lower accuracy for emotion identification in general, but non-significantly higher sensitivity for emotions in general.
	9 (49.3±9.0)	88.9	Manic	0	100?				
Vaskinn et al 2007	**21** (38.1±9.3)	47.6	NR but "euthymic"	NR	NR	**31** (30.7±9.6)	35.5	Face/Voice Emotion Identification and Discrimination Test	No differences in identification of facial or vocal emotion.

BDI, Bipolar I Disorder; BDII, Bipolar II disorder; HamD, Hamilton Depression Rating Scale; YMRS, Young Mania Rating Scale; MADRS, Montgomery-Asberg Depression Rating Scale

In recent years autism researchers have developed a theory of mind test which involves identifying an individual's mental state from a picture of the eye region of the face, the so-called 'Eyes Test' (Baron-Cohen et al., 2001). Rather than using pictures of stereotypical facial expressions, the stimuli are taken from everyday photographs in magazines and the popular media. The response options offered comprise more complex emotional terms than the traditional six basic emotions, with choices such as aghast, perplexed, bewildered and shy. Individuals with Asperger's Syndrome (a high functioning form of autism) perform poorly on this task, which has been attributed to a deficit in theory of mind skills.

The attraction of this test for individuals with bipolar disorder is the departure from the focus on basic emotions. The task involves interpreting expressions more akin to those likely to be encountered in everyday life. As such, performance may be more closely related to real-world functioning than for tasks using less ecologically valid stimuli. Additionally, the close tie between tasks grounded in discrete emotional models and a biological approach to understanding emotions has tended to frame the interpretation of impaired performance in terms of indicating an underlying problem in the neurocircuitry supporting emotion processing. However, the way in which the Eyes Test was designed taps more into a social constructivist approach (Kemper, 1987). The stimuli used were not posed specifically for the task and therefore had no inherent 'right' answer. The right answer was determined by asking 'healthy' individuals to spontaneously describe what they thought the person in the picture was thinking or feeling. The most frequent response generated was then taken to be the right answer, and offered as one of four possible response options for each pair of eyes in a multiple-choice question format. After further piloting, items producing the highest agreement among 'healthy' participants were selected for the test. Therefore, scores on the test as

a whole reflect the extent to which an individual shares a common understanding of these complex emotional terms.

One study so far has used this task in euthymic individuals with bipolar disorder (Bora et al., 2005) and patients showed significant impairment relative to controls.

IMPLICIT EMOTION TASKS

A number of different implicit emotion tasks have been used to explore emotion processing biases in bipolar disorder. One of the most commonly-used is the emotional stroop task (Bentall and Thompson, 1990, French et al., 1996, Kerr et al., 2005, Lex et al., 2008, Lyon et al., 1999, Malhi et al., 2005). All of the studies used a very similar Stroop paradigm including two emotion conditions (negative or depression-related words and positive or mania-related words) plus a control condition (Bentall and Thompson, 1990, French et al., 1996, Kerr et al., 2005, Lyon et al., 1999). One study was designed for an fMRI environment and used a push-button response rather than a verbal response (Malhi et al., 2005). Two of the studies used non-clinical samples (Bentall and Thompson, 1990, French et al., 1996). Participants in these two studies were selected on the basis of high scores on a hypomanic trait scale. Both of these non-clinical studies reported evidence of greater interference from negative words in the participants with high hypomanic traits, which was interpreted as evidence in line with the depression avoidance hypothesis. Of the studies in patients, three did not calculate interference scores (reaction time for emotional words minus reaction time for neutral words) (Kerr et al., 2003, Lex et al., 2008, Malhi et al., 2005). In euthymic patients, Malhi et al 2005 demonstrated slower reaction times for all words irrespective of valence, but did not examine the behavioural data for positive and negative words separately. Kerr et al (2005) reported general slowing of responses for all patient groups (manic patients, depressed patients, euthymic patients, and depressed patients with major depressive

disorder) compared to healthy controls across all stimulus types. Although no analysis of interference effects was conducted, examining the published data indicates that the manic patients showed the greatest difference between reading time for negative words compared to neutral words. The direction of the difference was consistent with greater interference from the negative words. Whether this difference was statistically significant would need to be tested. Lex et al (2008) reported no significant differences in reading times for negative, positive or neutral words. The interference effects were not examined directly but the mean reading times indicate that, relative to neutral words, patients took longer to name the negative words than the positive words. However, the difference is marginal and the variance relatively large making it unlikely that this difference would be statistically significant. Additionally, there were no differences between the groups in the valence of the words recalled from the Stroop task in a subsequent incidental recall task. The authors concluded that there was no evidence of cognitive processing biases on the included measures in their sample of euthymic patients with bipolar I disorder. The most comprehensive of the emotional Stroop studies (Lyon et al., 1999) included an explicit and an implicit self-esteem measure (a self-referent incidental recall paradigm) alongside the Stroop task. Manic patients scored highly on the explicit self-esteem measure and endorsed more positive interpersonal adjectives as true of self on the incidental task. However, the manic patients recalled more negative adjectives on incidental recall and showed greater interference from negative words on the Stroop task. For both of these latter tasks they performed similarly to a comparison group of depressed bipolar patients. This was taken as evidence in support of the depression avoidance hypothesis in that despite showing outward signs of mania, patients showed processing biases consistent with underlying depression.

Two studies have used an affective go/no go paradigm that involved pressing a button either in response to happy words or sad words in a continuous visual presentation of words of different valence (Murphy et al., 1999, Rubinsztein et al., 2006). One study reported no evidence of a negative bias in depressed patients with bipolar disorder (Rubinsztein et al., 2006), whereas the other found evidence of a negative bias (Murphy et al., 1999). One of the studies also included a group of manic patients, who showed a significant positive bias (Murphy et al., 1999). However it should also be noted that the manic patients in this study showed the slowest reaction times for the negative words compared to depressed patients or controls (although this difference did not reach statistical significance). This would be consistent with the greater interference from negative words reported in the emotional Stroop studies above.

Attentional cuing paradigms have also been modified for use as implicit emotion tasks in patients with mood disorders. Putman et al (2007) investigated attentional cueing in a student sample who scored highly on hypomanic traits (Putman et al., 2007). A gaze-cueing paradigm was used in which the eyes of a central facial image looked towards the right or the left to (validly or invalidly) cue a target. The facial expression of the central image was neutral, happy or fearful. The results showed an absence of an enhanced cueing effect for fearful faces in the group with high hypomanic traits that the authors interpreted to indicate reduced sensitivity to social cues of danger. Koster et al (2006) used a single-cue modified dot probe paradigm in a non-clinical sample to investigate the impact of anxiety and depression on engagement and disengagement with emotional faces (Koster et al., 2006). The paradigm involves presenting a facial image in one of two spatial locations which is either a valid or invalid cue for a subsequent target. Contrasting the response times taken to identify the target location for valid and invalid trials across different emotional expressions can be used to measure attentional engagement and disengagement with different facial emotions. In this study,

no effects of anxiety or depression on engagement or disengagement with angry, happy, or sad faces were evident. In contrast, enhanced engagement with angry faces was found in a similar study in a clinical sample of depressed patients with major depressive disorder (Leyman et al., 2007). In a different study, a group of depressed patients with bipolar disorder showed enhanced attention to angry faces (Leyman et al., 2009). The bipolar patients showed stronger cue-validity effects for angry faces than the control sample. Additionally they showed difficulties disengaging from both angry faces and positive facial expressions. The only study yet to use an emotional dot probe task in euthymic patients with bipolar disorder used a double-cue paradigm with emotional word pairs (Jongen et al., 2007). In this paradigm two words are presented either side of a central fixation and followed by a single target. The premise underlying the task is that response times to the target are faster if it appears in the same position as the word that had captured the participant's attention. By pairing positive or negative emotional words with neutral words it is theoretically possible to identify biases in attention to words of different valences. Euthymic patients showed a bias away from positive words. This pattern was also shown by a depressed sample of patients in the same study. Additionally the depressed patients showed a bias away from negative words.

SUMMARY AND ISSUES IN BIPOLAR DISORDER

Given the paucity of available data and variability in findings so far, understanding the implications of these results is difficult. The finding that manic patients tend to underestimate the intensity of sadness is consistent with a mood-congruent bias in affective processing. However, the lack of a compensating increase in the ratings of the intensity of happy faces could suggest the data are more in line with a defensive style where displays of negative emotion are devalued to reduce the level of threat they represent.

The data reviewed above suggest differences may be evident in the emotion recognition of patients with bipolar disorder, although these deficits are relatively subtle (i.e. there is not a global deficit) and there is a great deal of variability in the findings. Studies in euthymic patients have identified enhanced recognition of fear (Lembke and Ketter, 2002), enhanced recognition of disgust (Harmer et al., 2002), impaired recognition of fear (Venn et al., 2004), and no deficit in emotion recognition (Vaskinn et al., 2007). Differences in results stem from a number of factors relating to the sample used (e.g. bipolar I versus bipolar II, mood state at assessment) and methodology (e.g. emotion recognition versus emotion matching, still versus moving expressions, the use of different intensities of emotions). Until more studies use readily-comparable methodologies and patient samples it is difficult to discern whether there is evidence of a deficit (or enhancement) in emotion recognition.

Findings from the implicit measures suggest that some attentional biases may exist during mood episodes and that emotionally valenced stimuli may cause more interference or be harder to suppress when the emotional content of the stimulus is not relevant for the task. However, the evidence is mixed. Emotional Stroop paradigms have shown evidence in favour of greater interference from negative words in manic patients, which has been interpreted in line with the depression avoidance hypothesis. A different paradigm has been associated with a positive bias in manic patients, which is consistent with a more straightforward mood-congruent bias in emotion processing (Murphy et al., 1999). Attentional probe tasks using facial stimuli have indicated depression is associated with greater attentional engagement with angry faces (Leyman et al., 2009, Leyman et al., 2007). This could be interpreted in line with a broadly mood-congruent theme. Enhanced attention for anger in others may reinforce a sense that others are hostile or angry at you for failings or inadequacies. In euthymic patients there have been very few studies using implicit measures. One reported a general slowing on the emotional Stroop

paradigm, but did not analyse the interference effects for different classes of emotional stimuli (Kerr et al., 2005). Another reported a bias away from positive words in an attentional paradigm (Jongen et al., 2007). The same bias was evident in a sample of depressed patients in the same study.

It is clear that further studies are needed to clarify whether difficulties in emotion processing are associated with bipolar disorder and whether they can be informative about the nature of processing that takes place. In euthymic patients deficits or biases may represent a vulnerability to relapse whereas in acutely ill patients they may form part of a positive feedback loop that exacerbates a mood episode. Additionally there may be functional implications of emotion processing deficits. In patients with Schizophrenia deficits in emotion recognition have been associated with poorer functional outcomes (Kee et al., 2003). Developing an understanding of emotion processing in bipolar disorder and its relationship to social functioning could ultimately assist in developing new ways to improve functional outcomes for patients.

CHAPTER 4: INTRODUCTION TO PSYCHOSOCIAL FUNCTIONING IN BIPOLAR DISORDER

Impairment in functioning during acute episodes of illness in bipolar disorder is accepted as part of the diagnostic criteria for the disorder. Disturbances in mood and behaviour must cause "clinically significant distress or impairment in social, occupational, or other important areas of functioning" ((APA, 1994) p.358) in order to be of sufficient severity to merit a diagnostic label. However, when patients recover from an episode there is evidence that the level of functioning remains impaired in many cases (Chengappa et al., 2005, Coryell et al., 1993, Dion et al., 1988, Fagiolini et al., 2005, Gitlin et al., 1995, Keck et al., 1998, Tohen et al., 2000). There is notable variation in functional outcomes between patients, with one study in more than 3,500 patients identifying that after an acute manic episode, 28% of patients had low work impairment (as measured by a rating of 0 or 1 on a clinician-rated 5-point Likert scale that ranged from 0 (no work impairment) to 5 (unable to work due to mental illness)), 30% experienced moderate impairment (a score of 3 on the same scale), and 38% severe impairment (a rating of 4 or 5) (Goetz et al., 2007). This diversity in outcomes needs to be understood and developing a greater understanding of the factors that influence functional recovery may provide clues as to which strategies are best to support people with bipolar disorder to reach an optimal level of functioning between episodes.

WHAT IS "FUNCTION"?

In common-sense terms it is relatively easy to understand the concept of functioning – an individual's activities and capabilities in the social, occupational and recreational aspects of life all contribute to the person's functioning within society and

within their own social context. To pin down a specific technical definition is, however, far from straightforward. In an attempt to unify the language of functioning and aid the development of internationally comparable data relating to functioning, the World Health Organisation derived the International Classification of Functioning (ICF) which offers an idiographic and multi-faceted definition of functioning (Organisation, 2001). It is based around a biopsychosocial model of disability. In these terms, function is used very broadly to refer to something's operation, whether that be a body part or system, or something more abstract like a relationship or an activity. Impairments are a restriction or failure in the functioning of a body part or system (including psychological 'systems'). Whether that impairment leads to a difficulty in functioning (i.e. a difficulty in carrying out daily activities or participating in everyday life) depends partly on the level of the impairment itself, and partly on the personal and contextual factors specific to the individual. The ICF differentiates between functional performance – the level of function attained by the individual – and functional capacity – that which the individual is capable of in optimal circumstances. Restrictions to functioning may stem from individual factors (e.g. the severity of the impairment, the person's internal resources) and/or environmental factors (e.g. stigma leading to exclusion of affected individuals). The WHO's ultimate goal with the ICF is to collect more readily-comparable data about functioning and barriers to functioning in order to move an individual's functional performance as close as possible to their functional capacity and thus maximise their participation in everyday life.

In terms of bipolar disorder, conceptualising function in this way suggests there are several possible routes to improve functioning and it assists the understanding of the many and interacting influences on an individual's level of functioning. It also helps to identify what needs to be known in order to improve patients' functional outcomes. Research needs to develop a greater understanding of the factors that restrict

functioning in bipolar disorder. Factors leading to functional restrictions may be related to the severity of the impairment (which will be defined here as the (unknown) dysfunction of body structures or systems that 'causes' bipolar disorder) and/or to the absence of necessary environmental structures that would assist the individual in their daily lives (e.g. at the individual level, insufficient supportive resources such as a carer, or, at the societal level, stigma and insufficient accommodation for the difficulties faced by individuals with bipolar disorder). The ICF framework indicates that functioning could be improved by altering the individual factors (e.g. the severity of the impairment) and/or the environmental factors (e.g. providing more mental health workers, increasing awareness and reducing stigma). Altering the individual factors could mean minimizing the severity of the impairment by treating the illness in the most effective way possible and reducing bipolar disorder symptoms. Reducing the societal barriers is likely to include measures to increase awareness, education and acceptance of those with mental illness. Although vitally important in the battle to improve outcomes for patients with bipolar disorder, this latter aspect is beyond the scope of the present enquiry.

One important implication of the ICF framework indicates what has been shown in the literature so far, that reducing the level of impairment (which was defined above as the level of dysfunction in the relevant body structures or symptoms; reducing this impairment is assumed to be reflected in a lower level of mood symptoms, i.e. recovery from an acute episode) does not automatically improve an individual's functioning, since functional restrictions stem from more than the fundamental level of impairment experienced. The evidence that functional restrictions outlast syndromal recovery in a large proportion of patients may suggest that environmental barriers are still in place, or that the aspects of the illness most relevant to functional recovery are not currently being tackled by treatment interventions. There has been significant focus on residual symptoms and prior illness history as predictors of outcome in patients with bipolar

disorder, but a relative neglect of factors such as cognitive function and emotion processing – both of which have face validity as potentially significant influences on level of functioning. Both of these factors are discussed in more detail below (see page 83).

MEASUREMENT OF FUNCTIONING

The measurement and quantification of functioning is complex. There are numerous domains in which functioning can be measured (e.g. independent living, work, intimate relationships, friendships, relationships with family members, hobbies and interests) and within each domain level of functioning can be considered by how much (or whether) activity takes place in that domain, the quality of activity within each domain (e.g. is work performed to an acceptable standard, are relationships balanced and mutually satisfying), and whether the level of activity in that area is appropriate for that individual (e.g. in the occupational domain is the person capable of functioning to a higher level due to their level of education or experience). Owing to the subtleties and complexities of capturing all of these facets of functioning, all measures are necessarily imperfect in some way.

There are a vast array of instruments available (e.g. see (Andrews et al., 1994) which lists 20 scales for the assessment of functioning and 27 multidimensional scales that assess function alongside symptoms or quality of life) designed to capture functioning in a range of conditions across a variety of circumstances. Most measures capture functional performance, i.e. the level of functioning the individual attains, or the subjective degree of difficulty the individual gauges themselves to have/an observer gauges them to have (e.g. Specific Level of Functioning Scale (Schneider and Struening, 1983), Social Adjustment Scale (Weissman et al., 1978)). There are a smaller number of measures designed to capture functional capacity (e.g. UCSD Performance-Based Skills Assessment (Patterson et al., 2001a), Social Skills Performance Assessment (Patterson et

al., 2001b)), in part because such measures are more time consuming to administer and are difficult to validate. Whilst the *concept* of functional capacity is important, it remains a relatively abstract notion. Measuring someone's performance under optimal conditions is not always straightforward to achieve.

Another issue surrounding the measurement of functioning is the use of self-rated versus clinician-rated measures. Although any given individual is uniquely placed to gauge their level of psychosocial functioning, there are a number of factors that can cause objective and subjective ratings to diverge. Individuals with limited insight or self-reflection may find it difficult to judge their functioning in various domains in a manner that would match an external observer. By the same token, an external observer may lack important information, be unable to elicit the necessary information, or may be biased in their judgment, thus creating a picture that diverges from that seen through the individual's eyes. Functioning involves few, if any, absolutes and it is not possible to say which perspective is more valid than the other. There are instruments that incorporate information from more than one person (e.g. World Health Organisation Disability Assessment Schedule (de Jong et al., 1985)) and it seems any fruitful exploration of functioning requires several perspectives in order to be complete.

Measuring functioning in individuals with bipolar disorder has been particularly limited by the lack of suitable instruments available. Some of the measures commonly-used were originally derived for use in those with long-term severe mental illness, who may be institutionalised or have difficulty living independently in the community. Although some patients with bipolar disorder experience recurrent severe episodes which severely restrict even the most basic activities, most do not experience this scale of impairment and are able to live in the community and cater for their own basic needs. This may result in ceiling effects on those measures targeted at an inappropriate level.

Some measures have been derived for use in those with physical health conditions or physical disabilities and simply generalised across to psychiatric illness (e.g. the Medical Outcomes Survey Short Form 36 item (SF-36) (Ware and Sherbourne, 1992)). If they measure functional restrictions resulting from mental ill-health, the scales used can be very short and crude (e.g. the SF-36 contains 2 yes/no answer items addressing whether activity has been restricted due to mental ill-health) potentially restricting variance. Many measures confound symptoms and functioning, failing to provide a distinct measure of either (e.g. SF-36, Global Assessment of Function, Strauss-Carpenter Outcomes Scale). There are additional difficulties in measuring function that result from the nature of bipolar disorder itself. Functioning can vary markedly over the course of an individual's illness depending on the frequency and polarity of episodes. Individuals with bipolar II disorder, who experience hypomanic episodes but not manic episodes, may find that some components of their level of functioning improve during hypomanic episodes. The issue of when to measure functioning and which time point is most representative of the individual's general functioning is important in this patient group, yet difficult to ascertain.

In response to the need for scales suitable for assessing functioning in patients with mood disorders some authors have produced and validated rating scales for use in populations with mood disorders (Altshuler et al., 2002, Leon et al., 1999). The Life Functioning Questionnaire (LFQ) is a patient-rated measure that was designed to overcome the length and cumbersome nature of the Weissman's Social Adjustment Scale Self Report version (SAS-SR (Weissman et al., 1978)). The Longitudinal Interval Follow-up Evaluation Range of Impaired Functioning Tool (LIFE-RIFT) was designed as a brief clinician-rated measure of functioning for use in patients with mood disorders. Although emphasis was on its brevity in comparison with other measures that take more time to complete, a key consideration was also its suitability for use in populations with

mood disorder, whose level of impairment tends not to be as severe as those with a chronic and enduring psychotic illness.

As a final note with regard to measuring function, it is necessary to clarify that the emphasis in the present study is on function and not quality of life. Although a somewhat related concept, quality of life focuses more closely on the degree of satisfaction an individual derives from the activities they do and the relationships they have. Functioning relates more to whether or how often activity takes place in the various life domains. The decision to focus on functioning stemmed from its relatively more objective nature making it marginally easier to measure.

FUNCTIONING IN BIPOLAR DISORDER

Approximately a century ago, the major differentiating factor between the conditions we now refer to as Schizophrenia and Bipolar Disorder was the chronic unremitting course of the former versus the episodic nature of the latter which was associated with full recovery between episodes (Bentall, 2004). The ongoing attempt to improve our diagnostic systems and the progressive family of Diagnostic and Statistical Manuals produced by the American Psychiatric Association have seen gradual shifts in the boundaries for both Schizophrenia and Bipolar Disorder. One commentator suggested that the 'tightening up' of the Schizophrenia diagnostic criteria since DSM-III has served to reclassify some patients with poorer outcomes under the bipolar disorder label and has ultimately challenged the conclusion that patients who experience mania generally have a good outcome (Conus and McGorry, 2002).

Since the emergence of evidence that patients with bipolar disorder have poorer functional outcomes than was first believed (e.g. Conus et al., 2006, Tsuang et al., 1979),

there have been a number of studies investigating potential predictors of functioning in this patient group. What follows is a selective review of the evidence.

The present focus is on findings in patients who are euthymic or in the time period after an acute episode, seeing as functional impairment during an episode is only to be expected. However, several studies have poorly characterised samples or have not used specific criteria for euthymia and this must be held in mind when considering the findings.

As mentioned above, there is a large body of evidence that patients with bipolar disorder continue to exhibit functional impairment despite recovery from an acute episode (Chengappa et al., 2005, MacQueen et al., 2001). In one study, which followed-up 128 patients with bipolar disorder six months after admission for a manic episode, 77% of patients achieved syndromal recovery (i.e. they no longer met criteria for a current mood episode) whereas only 29% achieved functional recovery (i.e. reached the same level as the highest level achieved in the year before their episode) (Tohen et al., 2000). Considering only those patients who achieved syndromal recovery, almost 65% failed to achieve functional recovery. This study reported that different domains of functioning recover at different rates (Tohen et al., 2000), a finding that has also been reported by another study (Strakowski et al., 2000). In terms of residential functioning almost three-quarters had returned to pre-episode levels (the highest level achieved in the year prior to admission), whereas fewer than 40% met this definition of recovery in terms of occupational functioning (Tohen et al., 2000).

PREDICTORS OF FUNCTIONING

A large number of studies have investigated predictors of functional outcome, with the ultimate hope of identifying the key factors that contribute to poor outcomes in

order to understand where best to focus remediating interventions. A broad array of clinical and demographic variables have been tested and associated with poorer functional outcome. Most studies have examined factors relating to disease burden (e.g. number of episodes, residual symptoms, hospitalisations) and have reported associations. In a review, Conus & McGorry (2002) cited a number of studies that found a significant impact of interepisode symptoms on functioning, suggesting that subthreshold symptoms are a significant barrier to daily functioning (Conus and McGorry, 2002), a finding reported in a number of other studies (Dickerson et al., 2004, Fagiolini et al., 2005, Gitlin et al., 1995, Tohen et al., 2000). One study reported a significant association between polarity switches in the index episode and functional outcome 10 years later, noting that more than one polarity switch in the first episode was associated with poorer functioning at 10-year follow-up (Maj et al., 2002). However, this has not been investigated by any other studies. Age at illness onset has been investigated by several studies, but there are conflicting findings. Young age at onset has been associated with poorer functioning in some studies (Carlson et al., 2002, Strakowski et al., 2000, Tohen et al., 2000), but better outcomes in one (Tsai et al., 2001). The findings for comorbid substance misuse are also varied with one study reporting a significant negative impact of substance misuse on functional outcomes (Conus et al., 2006) and one reporting similar rates of functional impairment in patients with and without comorbid alcohol abuse (Tsai et al., 1997). Demographic factors have been reported as a significant predictor in at least three studies, with two reporting that being in a higher social class is associated with better functioning (Keck et al., 1998, Strakowski et al., 2000), and one reporting that receipt of disability pension and never having been married were both significantly associated with poorer overall functioning (Bauer et al., 2001).

COGNITION AND FUNCTIONING

There are a number of studies that have investigated the relationship between cognitive function and functional outcome (Altshuler et al., 2008, Atre-Vaidya et al., 1998, Dickerson et al., 2004, Dittmann et al., 2007, Laes et al., 2005, Malhi et al., 2007, Martinez-Aran et al., 2004a, Martinez-Aran et al., 2004b, Martinez-Aran et al., 2007, Simonsen et al., 2010, Zubieta et al., 2001). Direct comparison between studies is problematic as the different studies have used different functional outcome measures, different cognitive tests and different analyses. It is therefore relatively surprising that findings are broadly consistent – level of functioning is significantly correlated with aspects of cognitive function, most especially verbal memory.

Of the studies using a correlational approach (Atre-Vaidya et al., 1998, Dittmann et al., 2007, Malhi et al., 2007, Martinez-Aran et al., 2004a, Martinez-Aran et al., 2004b, Zubieta et al., 2001), only one reported no significant correlations between their measure of functioning (GAF) and any cognitive measure from a large battery of tests in euthymic patients (Malhi et al., 2007). However, this study only had a sample size of 10 and is therefore likely to have limited statistical power. A sample of 10 would require a Pearson correlation of $r>0.63$ to reach statistical significance. Unfortunately this study did not report the magnitude of the non-significant correlations to compare whether those found were of a similar magnitude to those reported in the other studies (in the region of $r=0.3$-0.5). Of the remaining correlational studies, two reported significant associations between GAF and both verbal list-learning tasks and various measures of executive functioning (Martinez-Aran et al., 2004a, Martinez-Aran et al., 2004b). However, these studies were conducted by the same group and it is unclear whether there was any overlap in the patient samples. Of the remaining three correlational studies, one divided the sample into service Veterans and a community-sample and used

different measures of functional outcome for the two samples (Atre-Vaidya et al., 1998). In the Veteran sample, scores on the Impairment Rating Scale were associated with list-learning such that poorer functioning was associated with worse verbal memory (as measured by several indices of the California Verbal Learning Test (CVLT), namely total recall, short-delay free recall and long-delay free recall). In the community sample, scores on the clinician-rated Structured and Scaled Interview for Maladjustment correlated negatively with verbal recall (again total recall, short-delay free recall and long-delay free recall on the CVLT) and additionally correlated negatively with one aspect of executive function (verbal fluency). One further study study reported a significant association between ratings on the Social and Occupational Functional Assessment Scale and both immediate verbal paired associates learning (from the Wechsler Memory Scale) and scores on the Stroop colour-word test. The final study reported a significant relationship between a self-rated functioning measure (the Social Adjustment Scale) and performance on executive and psychomotor measures (Trail Making Test part B and the Letter Number Sequencing Test). In general the findings indicate significant associations between functioning and verbal memory, psychomotor speed, and various aspects of executive function. Of the three studies that did not report either any associations or no association with verbal memory, one had a small sample size and therefore low statistical power (Malhi et al., 2007), and the remaining two did not assess verbal memory via a standard list-learning or paragraph recall paradigm (Dittmann et al., 2007, Zubieta et al., 2001).

A single study used a different approach and divided a euthymic clinical sample according to good or poor functioning as measured by the GAF then compared the two groups on cognitive functioning (Altshuler et al., 2008). Despite a small sample size (n=7 in each group), the poor functioning group showed statistically worse verbal memory (as assessed with the California Verbal Learning Test) and executive function (as measured

with the Wisconsin Card Sorting Test) than the good functioning group. Additionally, the group with poor functioning performed significantly worse than healthy controls, whereas the good functioning group showed no differences from controls.

Correlational studies are an important first step in identifying factors likely to play a role in functional impairment in patients with bipolar disorder, they are limited by the fact that several predictors cannot be considered simultaneously. As there are undoubtedly a large number of factors that influence functioning, considering several factors at the same time (such as in regression analysis) allows an understanding of the relative strength and contribution of different explanatory variables. Of particular interest is the role that cognitive function plays in predicting function as compared with other factors such as symptoms or illness history variables. Four studies have employed regression analysis in this way in patients with bipolar disorder including at least one measure of cognitive function alongside other predictors (Dickerson et al., 2004, Laes and Sponheim, 2006, Martinez-Aran et al., 2007, Simonsen et al., 2010). However, only one of these has been conducted using a euthymic sample (Martinez-Aran et al., 2007).

The study in euthymic patients used the GAF to rate the functioning of 75 euthymic bipolar patients and, from a selection of clinical and cognitive predictors, entered those that correlated significantly with GAF ratings into a stepwise regression model (Martinez-Aran et al., 2007). The final regression model included just two predictors – delayed verbal memory (CVLT long delay free recall) and number of medications. The model had modest explanatory power, explaining 22% of the variance in GAF ratings. The authors examined a second model including clinical predictors reported as significant predictors in previous studies (subclinical symptoms, length of illness, number of episodes, and number of hospitalizations) as well as the cognitive predictors. Once again, the strongest single predictor was delayed verbal memory,

accounting for 21% of the variance in GAF ratings. No other variables reached statistical significance.

The other three studies that explored the relationship between functioning and cognition all reported that symptom measures were also significant or were the primary predictors of functioning (i.e. explained the largest proportion of variance). Dickerson et al (2004) divided their sample of 117 bipolar patients into three groups according to their work status – no work activity, part time or voluntary work, or full time competitive work. Using logistic regression with work status as the dependent variable, they examined the association between work status, clinical factors, demographic factors and cognitive factors. They reported a significant association between work status and cognitive function (total score on the Repeatable Battery for the Assessment of Neuropsychological Status), current symptoms (score on the Brief Psychiatric Rating Scale), a clinical factor (the number of previous hospitalizations), and a demographic factor (maternal education). Variables that did not enter the model included other measures of symptoms (level of manic symptoms), medication usage, and scores on other cognitive tests. The explanatory power of the model as a whole reached statistical significance and the variables identified successfully predicted the employment status of 64% of the sample (Dickerson et al., 2004). Laes & Sponheim (2006) reported in a sample of 27 bipolar patients that scores on a self-report scale, the Social Adjustment Scale, correlated significantly with an executive measure (excess moves on the Tower of London test). However, when entered into multiple regression analysis alongside clinical variables, only the level of psychotic symptoms entered the model. The overall explanatory power of the model was poor with an R^2 of only 10% (Laes and Sponheim, 2006). Simonsen et al (2010) investigated clinician- and self-rated function in two different groups of patients with bipolar disorder. One group had previous history of psychotic symptoms (n=64), and the other had no prior history of psychotic symptoms

(n=56). They reported that clinician-rated function and self-rated function were associated with similar predictors. For both groups, current symptomatology explained more variance in functioning than neurocognitive function. For the group with psychotic bipolar disorder, cognitive indices were not a significant predictor. For the group with non-psychotic bipolar disorder, neurocognitive variables were significant predictors when entered alone (i.e. without symptom measures), but were reduced to statistical non-significance when symptom measures were added to the model. The explanatory power of the models ranged from 24%-50%.

DISCUSSION

The findings with regard to predictors of functional outcome in bipolar disorder are best described as varied. There is relatively little direct contradiction, but there is a lot of variation stemming from the fact that studies have tended to measure different potential predictors and use different methodologies (both in terms of different measures of functioning and different approaches to statistical analyses). Although a number of studies have reported significant associations between various clinical variables and functioning, these studies have tended not to include measures of anything else, or have used a purely correlational approach, and therefore have been unable to ascertain the relative contribution of different predictors. What has been measured sets an obvious limit on both the relationships that can be found and the conclusions that can be drawn about the relative importance of variables from different domains. Of the studies that included both clinical and cognitive predictors in regression analysis, the study in euthymic patients reported that cognitive predictors explain more variance than clinical factors to the extent that clinical factors no longer entered the regression model once cognitive factors had been accounted for. The three studies in patients in different clinical states reported either that level of symptoms was the only predictor of

functioning, or that symptom-level explained more variance than cognitive function. It is understandable that within an episode, level of symptoms is a key determinant of level of functioning and therefore in mixed clinical samples the greater variation in symptom status is more likely to show a relationship with functioning. The data from one study indicates that when patients are euthymic, cognitive factors become primary.

There are a number of limitations of studies conducted to date. Firstly and most importantly the functioning measures used have often incorporated a measure of symptoms and thus confounded symptoms and functioning. It is therefore unclear whether any significant associations relate solely to functioning. For example, the GAF, Strauss Carpenter Outcome Scale and the Impairment Rating Scale all measure both symptoms and level of daily activities. Although three studies described using the GAF for functional ratings only (Altshuler et al., 2008, Martinez-Aran et al., 2007, Simonsen et al., 2010), none of the others described making any accommodation for this confound. In these circumstances it is impossible to conclude whether predictor variables are associated with functioning, symptoms or both.

The reported findings have demonstrated weak or moderate relationships between functioning and the variables investigated. Regression analyses have generally explained less than 30% of the variance in function and correlation coefficients rarely exceed 0.5 (25% of variance explained), suggesting that there are still important explanatory variables missing from the analyses conducted to date. One potentially important factor – social information processing – has yet to be examined as a predictor of functioning in people with bipolar disorder. In people with Schizophrenia, there is evidence that worse recognition of facial and vocal emotion is associated with poorer work functioning and ability to live independently (Kee et al., 2003). Examining a similar

relationship in bipolar disorder may account for some proportion of the remaining variance in functioning.

Related to this point is the fact that studies have tended to consider a small number of domains, but sample highly from those domains (e.g. clinical or demographic predictors, but with multiple measures of each). Items within a single domain are likely to share a great deal of common variance and unless the aim is to identify which from a number of measures is the one that captures the most common variance (and thus save time in future studies by simply employing that single measure) then this approach is likely to give a misleading impression that a single test index, illness variable, or demographic factor carries important weight in explaining variation in functioning between individuals over and above other similar measures. Social functioning and cognitive functioning are both complex and multi-faceted constructs. It seems to be an unrealistic goal to identify precise predictions of functioning by a small and limited number of specific variables. In these early stages of enquiry, an approach is needed that accepts that this goal is unrealistic and that initial analyses may identify candidate predictors, but these candidates should be taken as representative of a pool of potential predictors that all overlap in the variance they explain. Subsequent analyses can use data reduction techniques to hone in on this common variance (e.g. factor analysis) and use a summary of it (e.g. factor scores) in place of the individual specific predictors thereby providing an idea of the domains of most importance for psychosocial function. This approach is greatly enhanced if measures from several different domains are examined simultaneously.

Furthermore, with respect to identifying the key cognitive areas that relate to functioning in patients with bipolar disorder, it is helpful to consider briefly the results of similar investigations in patients with Schizophrenia. One review reported that

performance on tests of memory, executive function and vigilance has shown a relationship with social functioning (Green, 1996). One group investigated the mechanism by which neuropsychological function may impact on social functioning (Bowie et al., 2006). They reported that functional capacity (the behaviours and abilities an individual is capable of under optimal conditions) mediates the relationship between neuropsychological function and functional performance (the behaviours and abilities that are actually produced or used in the real world). In other words, neuropsychological function seems to act as a limit-setter, restricting the highest level of functioning an individual could achieve, but other factors (such as negative symptoms) then go on to impact how that functional potential is translated into actual functioning. One author postulated that the key cognitive skill that may underlie the relationship between cognitive function and social functioning is learning potential (Green et al., 2000). Loosely defined, learning potential shifts the focus from knowledge an individual already has, to the amount of information they are capable of learning. In Green's formulation, learning potential is seen as a potential mediator between neurocognition and skills acquisition, and skills acquisition is seen as an important determinant of functioning. Investigating learning capacity in patients with bipolar disorder and examining the relationship with social functioning would identify whether it is a construct of interest and relevance for the nature of cognitive impairment in bipolar disorder and also the way in which it may impact on social functioning.

To date, the small number of studies that investigated the relationship between cognitive function and psychosocial function have predominantly used clinician-rated functional assessments. Incorporating self-rated measures may be enlightening in order to identify whether differences in perspective relate differently to potential predictors.

In summary, to develop our understanding of psychosocial functioning impairment in bipolar disorder, a comprehensive assessment of functioning from both a clinician- and self-rated perspective is needed so as to capture its many shades and subtleties to maximum effect. To understand potential predictors of functioning, several areas need to be examined simultaneously in the hope of maximizing the variance explained.

CHAPTER 5: NEUROPSYCHOLOGICAL FUNCTION IN BIPOLAR DISORDER

INTRODUCTION

As discussed in Chapter 2, there is a large body of evidence of cognitive dysfunction in bipolar disorder. Meta-analyses have demonstrated moderate to large impairments in various aspects of executive function, verbal memory and visual memory in euthymic patients (Arts et al., 2008, Bora et al., 2009, Kurtz and Gerraty, 2009, Robinson et al., 2006, Torres et al., 2007). The aim of the present chapter is to take a broad look at cognitive function in a further sample of euthymic patients with bipolar disorder, with future chapters going on to explore different aspects of cognition in greater depth.

Much of the work to date has focused on executive and verbal memory deficits as the areas which show the greatest degree of impairment. The battery for the present study was therefore selected to draw heavily on these areas.

Far from being a unitary construct, 'executive function' encapsulates a broad range of cognitive processes. Debate continues as to whether executive function is a single entity that performs multiple functions, or a set of separate control processes that operate relatively independently, with some evidence in favour of the latter (Baddeley, 1996, Miyake et al., 2000). As such, describing clinical populations as demonstrating 'executive impairment' requires further qualification as to which aspects of executive function are impaired. Several authors have proposed taxonomies of executive processes, and the one employed presently was suggested by Baddeley (Baddeley,

1996). This is a longstanding and highly influential model, aspects of which have been used in similar studies from our own research group. Baddeley identifies four subsets of executive functions, specifically: set-shifting, generational fluency, response inhibition, and mental manipulation. Tests designed to measure each of these aspects of executive function are included, as well as tasks assessing one aspect omitted from Baddeley's formulation – planning. Tests of planning have been included as impairment in this aspect of executive function has been reported in patients with bipolar disorder (Thompson et al., 2005).

To provide a thorough assessment of verbal memory function, the present study uses both list-learning and passage-recall paradigms. Much of the work investigating verbal memory performance in bipolar disorder has used list-learning tasks. Learning an unconnected series of words is perhaps a good parallel to acquiring fact-based knowledge, but – on the face of it – shares less in common with remembering narratives or prose, a task which is encountered on a daily basis. The incorporation of a semantic structure and a 'gist' (a fundamental meaning that can be expressed in a number of different ways) in passage recall tests that is usually lacking (or at least obscured) in list-learning tasks allows further exploration of the nature of the verbal memory deficit in bipolar patients.

As one ultimate aim of the current investigation is to derive a better understanding of the relationship between cognitive function and psychosocial function in patients with bipolar disorder, assessing cognitive constructs that have been shown in previous studies of bipolar disorder or other psychiatric illnesses to be of relevance for social functioning is important (for a more in-depth discussion see Chapter 9). In patients with Schizophrenia, performance on tests of memory, executive function and vigilance has shown a relationship with social functioning (Green, 1996). One group investigated

the mechanism by which neuropsychological function may impact on social functioning (Bowie et al., 2006). They reported that functional capacity (the behaviours and abilities an individual is capable of under optimal conditions) mediates the relationship between neuropsychological function and functional performance (the behaviours and abilities that are actually produced or used in the real world). In other words, neuropsychological function seems to act as a limit-setter, restricting the highest level of functioning an individual could achieve, but other factors (such as negative symptoms) then go on to impact how that functional potential is translated into actual functioning. One author postulated that the key cognitive skill that may underlie the relationship between cognitive function and social functioning is learning potential (Green et al., 2000). Loosely defined, learning potential shifts the focus from knowledge an individual already has, to the amount of information they are capable of learning. In Green's formulation, learning potential is seen as a potential mediator between neurocognition and skills acquisition, and skills acquisition is seen as an important determinant of functioning. Learning potential (or capacity) is a difficult construct to measure. Green (2000) suggests that the relationship between verbal learning and social functioning emerges so consistently because verbal learning tests are the closest measure of learning potential used in most test batteries. They involve dynamic assessment, often with multiple presentations of the same material, both of which are important components of a learning potential measure. However, verbal list learning tasks are not a perfect measure and are clearly not designed for this purpose. To investigate the concept of learning potential in patients with bipolar disorder, a different measure is used – repeat administration of the digit symbol substitution test (DSST, Corporation, 1997)). The task involves pairing numbers with a non-numeric symbol as quickly as possible in a given time limit. Assessing improvement over repeated administrations of the same task version gives a measure of the extent to which an individual can benefit from incidental

learning and therefore potentially reflects their learning capacity. In patients who had throat surgery, this measure was found to correlate strongly with outcome of post-surgery rehabilitation, in that patients with a higher learning capacity required less intensive speech therapy input (Ho et al., 2005).

METHODS

PARTICIPANTS - PATIENTS WITH BIPOLAR DISORDER

Individuals with bipolar disorder were recruited from secondary and tertiary care services via their treating clinician. Participants underwent an initial screening session conducted by a psychiatrist which involved confirming diagnosis using the mood disorders section of the SCID-IV (First et al., 1995), assessment of current depressive and manic symptoms with the Hamilton Depression Rating Scale (HDRS; Hamilton, 1960) and Young Mania Rating Scale (YMRS; Young et al., 1978) respectively, assessment of comorbidities using the Mini-International Neuropsychiatric Interview (Sheehan et al., 1998), and a brief patient-reported illness history. The assessing psychiatrist was trained to at least Senior House Officer level (Foundation Year 3), but not all patients were assessed by the same psychiatrist and inter-rater reliability on the clinical ratings was not established. Patients subsequently completed self-rated mood, anxiety and self-esteem measures. Additionally, at the screening session all participants were tested for colour-anomalous vision with the Neitz Test (Neitz et al.). Inclusion criteria for the study comprised: 1) DSM-IV diagnosis of bipolar disorder; 2) currently euthymic (scores of <8 on the HDRS and <8 on the YMRS); 3) aged between 18 and 65. Exclusion criteria comprised 1) current alcohol misuse or dependence; 2) history of head injury with loss of consciousness lasting more than 5 minutes; 3) known neurological illness or relevant

major medical illness; 4) ECT within the last 6 months; 5) learning disability or difficulty with fluent use of the English language. Patients were not excluded for comorbid anxiety disorders or for use of psychotropic medication. After the screening session, patients completed weekly mood ratings of depressive and manic symptoms (using the Beck Depression Inventory (BDI) (Beck et al., 1961) and the Altman Mania Rating Scale (AMRS) (Altman et al., 1997) respectively) for four weeks, then returned and had their mood symptoms rated by a psychiatrist once again with the HDRS and YMRS. Participants continuing to score below the designated cut-offs at the second assessment were then tested with a battery of neuropsychological tests. Shortly before the tests took place, participants were rated on three clinician-rated psychosocial functioning measures (see Chapter 9 for full details of the measures) and rated themselves on three self-rated social functioning measures (see Chapter 9). Participants scoring outside of the designated range on the clinician-rated mood scales at the second assessment were monitored until they were euthymic for at least four weeks, at which point they completed the social functioning measures and the neuropsychological tests.

In all, 57 patients with bipolar disorder underwent screening. Thirty three patients successfully completed all study measures. A further six had incomplete data sets owing to a range of issues including hardware failure and insufficient time to complete the assessment (e.g. two participants had travelled a long way in order to participate and owing to factors outside of the participant's or the experimenter's control were unable to begin the assessment on time, and were therefore unable to complete the testing session before having to leave to travel home). In these instances, as there was no reason to believe the non-completion of measures was likely to introduce bias, the data for these participants was included where it was available. Of the 18 patients who did not proceed beyond screening, two did not meet full criteria for bipolar disorder, eight had levels of depression above the cut-off that did not meet

criteria for euthymia within the study period, one had levels of manic symptoms that exceeded the cut-off and did not meet criteria for euthymia within the study period, one participant had comorbid alcohol misuse, one exceeded the upper age limit, one had an autistic-spectrum disorder, and four withdrew from the study. A flow chart detailing this information is provided in Appendix 2 on page 311.

PARTICIPANTS – CONTROLS

Healthy controls with no personal history of psychiatric illness and no family history of bipolar disorder were recruited by advert and word of mouth from the North East of England. Controls underwent an initial screening session conducted by a psychiatrist which involved administering the MINI to confirm no current psychiatric diagnosis, ascertaining any personal or family history of psychiatric illness, and assessment of current depressive and manic symptoms with the HDRS, YMRS and BDI. Participants meeting inclusion criteria then completed the same self-ratings and colour-vision assessment as the patients, and additionally completed the objective and subjective social functioning measures. Inclusion criteria comprised: 1) aged between 18 and 65; 2) score on the BDI of <9. Exclusion criteria comprised: 1) personal history of psychiatric illness that required treatment with antidepressants or psychological therapy; 2) family history of bipolar disorder in a first degree relative; 3) history of head injury with loss of consciousness lasting more than five minutes; 4) known neurological illness or relevant major medical illness; 5) learning disability or difficulty with fluent use of the English language. After the screening session, a second session was arranged for neuropsychological testing at a time convenient to the participant (as there was no need to monitor for euthymia, control participants were not subject to the same four-week delay between screening and testing).

Control participants were group-matched as closely as possible to the patient group on age, gender, and years of education. Control participants were paid for their participation.

In total, 31 individuals underwent screening and twenty eight of those successfully completed the study. Of the remaining three, two had previous history of psychiatric illness and one withdrew from the study. A flow chart detailing this information is provided in Appendix 2 on page 311.

ETHICS

The study was approved by the Northumberland Local Research Ethics Committee. All participants gave written informed consent to participate after the study aims and procedure were explained and before the screening session.

MEASURES – SELF-RATED QUESTIONNAIRES

Beck Depression Inventory (Beck et al., 1961) – This is a 21-item scale assessing depressive symptoms over the preceding week. Each item is rated from 0-3, giving a minimum possible score of 0 and a maximum possible score of 63. Higher scores indicate worse depression. The standard cut-offs are: 0-9, no depression; 10-18, mild/moderate depression; 19-29, moderate/severe depression; 30 or above, severe depression.

Altman Mania Rating Scale (Altman et al., 1997) – This is a 5-item scale assessing manic and hypomanic symptoms over the preceding week. Each item is rated from 0-4, giving a minimum possible score of 0 and a maximum possible score of 20. Higher scores indicate more severe manic symptoms. The standard cut-off for likely hypo/mania is six or above.

Rosenberg Self-Esteem Scale (Rosenberg, 1965) – This is a 10-item scale measuring self-esteem. Each item is a statement indicating how someone may feel about

themselves (e.g. “On the whole I am satisfied with myself” “At times, I think I am no good at all”) and the participant rates the extent to which they agree with each statement according to a 4-point scale (strongly agree, agree, disagree, strongly disagree). Participants are asked to rate the statements according to how they generally feel about themselves. Five of the statements are positive self-statements and five are negative. Positive items are scored from 0-3, with 0 representing ‘strongly disagree’ and 3 ‘strongly agree’. Negative items are reverse scored such that 0 represents strongly agree and 3 represents strongly disagree. The scores for each item are added together yielding a total between 0-30 with higher scores representing higher levels of self-esteem. To minimise the introduction of bias caused by using the term ‘self-esteem scale’ on the questionnaire itself, it was presented to participants as the ‘Rosenberg Attitudes to Self Scale’.

Dysfunctional Attitudes Scale (Power et al., 1994) – This is a 24-item scale measuring attitudes in three domains – achievement, dependency, and self-control. Each item states a belief or attitude and the participant rates how much they agree with the statement on a seven-point scale (totally agree, agree very much, agree slightly, neutral, disagree slightly, disagree very much, totally disagree). Participants are asked to rate each item according to how they think most of the time. The statements can be divided into three subscales – achievement-related items (e.g. ‘People will probably think less of me if I make a mistake’ ‘My life is wasted unless I am a success’), dependency-related items (e.g. ‘If others dislike you, you cannot be happy’ ‘I am nothing if a person I love doesn’t love me’), and self-control-related items (e.g. ‘I ought to be able to solve my problems quickly and without a great deal of effort’ ‘A person should be able to control what happens to him’). Items are scored from 1-7, with a score of 1 equating to ‘totally disagree’ and a score of 7 indicating ‘totally agree’. Four items are reversed-scored (three from the dependency subscale and one from the self-control subscale). The

minimum total score is 24 and the maximum is 168. There are eight items for each subscale (yielding minimum scores of 8 and maximum scores of 56 for each). Higher scores indicate more dysfunctional beliefs. The items are presented as an achievement-item followed by a dependency-item followed by a self-control item all throughout the questionnaire.

State-Trait Anxiety Inventory (Spielberger and Gorsuch, 1970) – This is a 40-item scale divided into two equal parts, the state anxiety inventory and the trait anxiety inventory. The state inventory lists 20 'I feel…' or 'I am…' statements (e.g. 'I feel calm' 'I am tense') and the participant indicates the extent to which each statement is true for them at that moment in time on a four-point scale (not at all, somewhat, moderately so, very much so). The scale is scored from not at all = 1 through to very much so = 4, and ten of the items are reverse scored. The total score equals the sum of the scores for each item and ranges from 20-80, with higher scores indicating higher 'in the moment' anxiety.

The trait inventory similarly lists 20 statements relating to an individual's general level of anxiety (e.g. 'I worry too much over something that really doesn't matter' 'I make decisions easily') and the same four-point scale is used to rate how much the statement applies generally. Scoring is as for the state anxiety inventory with higher scores indicating higher general or trait levels of anxiety.

MEASURES – NEUROPSYCHOLOGICAL TESTS

1. **DIGIT SYMBOL SUBSTITUTION TEST (WECHSLER ADULT INTELLIGENCE SCALE III (WAIS-III) (CORPORATION, 1997))**

In this task the participant is presented with the digits 1-9 at the top of the page with a different symbol underneath each. Below is a grid showing the symbols with blank boxes underneath. The participant is instructed to put the correct digit in the box under each symbol using the guide at the top as a key. The test has seven rows of symbols with 20 in each row (the first seven items are for practice). The participant is timed and told to complete the task as quickly as possible without making any mistakes and without missing any items out. The administration time was 90 seconds. The outcome measure is the number of items correctly completed in 90 seconds.

To measure participants' incidental capacity to learn the digit-symbol pairings, this test was administered on four separate occasions throughout the testing session. Exactly the same version of the task was used on each occasion. The raw scores were recorded, and the percentage change between each consecutive pair of trials and the linear slope of performance change over the four administrations were calculated.

2. VERBAL FLUENCY – PHONETIC (LEZAK ET AL., 2004)

The participant is given a letter of the alphabet and timed for 60 seconds to say aloud as many words as they can think of that begin with that letter, subject to certain constraints (no proper nouns, no numbers, no repeats, and not using the same word again with a different ending). The task is completed three times with three different letters: F, A and S. Outcome measures comprise the total number of appropriate words generated, the total number of repeats, and the total number of words that break the rules.

3. VERBAL FLUENCY – SEMANTIC (LEZAK ET AL., 2004)

Similar to the above task, the participant is given a semantic category and timed for 60 seconds to say aloud as many words as they can think of that belong in that

category, avoiding repeats. The task is completed three times with three different categories: animals, fruits & vegetables, and occupations. Outcome measures comprise the total number of appropriate words generated, and the total number of repeats.

4. HAYLING SENTENCE COMPLETION TEST (BURGESS AND SHALLICE, 1997)

This task has two parts. In part 1 the participant is read a series of 15 sentences each with the last word missing and is asked after each one to complete the sentence as quickly as possible with a word that makes sense (e.g. "He posted a letter without a …" "Stamp"). For part 2 of the task the participant is read a second (different) series of 15 sentences, again with the last word missing. They have to complete each one as quickly as possible with a word which does not fit at the end of the sentence (e.g. "The whole town came to hear the mayor …" "Yellow"). The challenge in part 2 is to suppress the more obvious sensible completion of the sentence and think of an entirely unrelated word. In both parts the participant's responses are timed and rounded down to the nearest second.

There are several outcome measures for this task. For part 1 the sum total of the time taken for each item (in seconds) is calculated and scaled between 1 and 7 as per the standard scoring guidelines (higher scaled scores indicate quicker responses). For part 2 total time taken is calculated and scaled in the same way – including all items irrespective of errors. For part 2, two different types of error are scored – category A and category B. Category A errors involve sensible completion, i.e. the offered word makes sense at the end of the sentence and fits with the context of the sentence (e.g. "None of the books made any…" "Sense"). Category A errors are converted to an A score between 3 and 78 according to the scoring guide. The A score increases non-linearly with number of A errors made, reflecting the fact that making several A errors is proportionately worse (or rarer) than only making a few. Category B errors involve completion of the

sentence with a word which is somewhat connected to the meaning of the sentence (or some part of it), but is not a direct completion (e.g. “None of the books made any…” “Difference”). There are many possible types of category B errors, but all indicate that the response given has some connection with the sentence rather than being totally unrelated, and thus the participant has not fully inhibited sentence-meaning. Category B errors are converted to a B score ranging between 1 and 50 using a non-linear transformation. Summing together the scaled times and error scores gives a total scaled score, which in turn is converted to an overall scaled score ranging from 1-8, with a higher score indicating better performance.

5. REY AUDITORY VERBAL LEARNING TEST (RAVLT) (REY, 1964)

In this task the participant is read a 15-item word list and is asked to recall as many items from the list as they can (trial 1). The same procedure is followed four further times, with the same list of words (list A) being read to the participant and recall measured after each presentation of the list (trials 2-5). After the fifth presentation, a different 15-item word list, list B, is read once to the participant and recall is assessed. After this distractor list, the participant is asked to recall as many items as possible from list A without hearing the words again (trial 6). After a 30minute delay (in this instance filled with non-verbal tasks), recall of list A is assessed again (trial 7), followed by a recognition trial. In the recognition trial, participants see 50 words on a page – all 15 from list A, all 15 from list B, and 20 words that were not on either list – and are asked for each word to say whether it was on the first list, the second list, or neither list.

There are a number of performance indices derived from this task which reflect various aspects of learning and memory: immediate recall/memory span is assessed by trial 1 of list A; total learning is assessed by total words recalled for trials 1-5 of list A; susceptibility to interference is measured by trial 6 of list A; delayed recall is measured

by trial 7 of list A; forgetting is measured by the percentage retained between trials 6 and 7; and recognition memory is assessed by the number of hits on the recognition trial.

In the standard administration of the task, list A is presented in the same order on trails 1-5. In the present study, the word list was presented in a random order on each presentation (see Chapter 6 for more detail).

6. Wisconsin Card Sorting Test – 128 card version (Heaton et al., 1993)

For the details and results of this task, see Chapter 7.

7. Computerised version of the Abstract Designs Self-Ordered Pointing Test (Petrides and Milner, 1982)

Presented on a computer with a touch screen monitor, for the first level the participant sees an array of four abstract shapes and is instructed to touch each of the shapes only once in any order. Each time the participant touches one of the shapes, the array is rearranged on the screen. For successful performance the participant must remember which shapes they have already touched in order not to touch them again. There are four levels of the task – 4 shapes, 6 shapes, 8 shapes and 10 shapes. There are three trials at each level (using exactly the same shapes for each trial). Each level uses a different set of shapes.

The outcome measures comprise the number of errors made at each level of the task, the highest level at which the participant managed a correct response (i.e. one trial of unique touches), and the maximum memory span at each level (the number of correct touches made before an error).

8. Trail Making Test – A & B (Reitan, 1958)

Part A of the trail making test presents the participant with the numbers 1 to 25 on an A4 page the participant is instructed to connect the numbers in increasing numerical order as quickly as possible using a single continuous line. Total time to completion is recorded, with errors simply reflected in a slowed time (participants are taken back to the point they made the error and proceed from there).

In part B the participant is presented with a mixture of numbers and letters on a page and instructed to connect the numbers in numerical order and the letters in alphabetical order using a single continuous line as quickly as possible, but to alternate between numbers and letters. Total time to completion is recorded.

To assess set-shifting controlling for basic psychomotor speed, switch time is calculated (time for part B minus time for part A).

9. National Adult Reading Test (Nelson, 1982)

This test is used to estimate premorbid IQ. The participant is asked to read aloud a list of 50 irregularly-spelled words. The number of errors is converted into an estimated premorbid full-scale IQ score.

10. Logical Passages Test (Wechsler Memory Scale III) (Wechsler, 1997)

This test of verbal memory involves reading a supra-span passage of text to the participant, presented as a story, and assessing immediate recall. Two different stories are presented one after the other with recall assessed after each. Delayed recall was not assessed. The outcome measures comprise the summed total of story units correctly recalled for both stories and the summed total of story themes correctly recalled.

11. Simultaneous and Delayed Match to Sample

This task involves presenting the participant with an abstract shape in the centre of the screen. In the simultaneous condition, four shapes are presented below the central shape and the participant must pick out which of the four matches the central shape by touching the relevant option on the screen. In the immediate condition, the central shape is masked with a solid rectangle and the four options appear underneath without any interval. In the delayed condition, there is a four-second delay between the central shape being masked and the response options appearing.

To avoid ceiling effects in healthy participants, a version of this task was designed specifically for this study. Using a set of shapes derived using the principles described by Attneave & Arnoult (1956), 36 meaningless shapes were selected – 6 with each of 4-, 6-, 8-, 12-, 16-, and 24-points (Attneave and Arnoult, 1956). The shapes were selected to be abstract and have low nameability to minimise the use of straightforward verbal labeling strategies.

The task involves 36 trials in total. One-third of trials (n=12) are simultaneous trials, one-third are 0-second delay (i.e. immediate) trials and one-third are 4-second delay trials. Of the 36 trials, 18 were designated as 'easy' trials and 18 as 'hard'. In the easy trials, the three distractor shapes were selected to bear less similarity to the target shape than in the hard trials. Specifically, for each easy trial, two of the distractors were different shapes with the same number of points as the target and one was a shape with a different number of points to the target (see Figure 5.1A). By contrast, for hard trials, the distractors were all variants of the target shape (see Figure 5.1B). The task was piloted in a small number of individuals who were not due to participate in the study. Results indicated the hard trials of the task did not result in ceiling effects in these individuals.

The 36 trials were presented in the same random order for each participant with a break half way through. Trial presentation was counterbalanced such that both halves of the task contained an equal number of hard and easy trials, an equal number of each delay type, and an equal number of target shapes with 4-, 6-, 8-, 12-, 16-, or 24-points. In the two halves of the task, the target shape appeared as evenly as possible in each of the four response locations (it was not possible to balance this perfectly as the number of trials in half the task was not divisible by four, but this was balanced across the task as a whole).

In the 0-second and 4-second delay trials the target shape was displayed for four seconds then covered by a solid mask. In all trials the trial was ended once the participant gave their response. Participants were not instructed to respond as quickly as possible. The task was presented using Superlab 4.0 stimulus display software (Cedrus) and responses were recorded using a 15" CTX resistive touch-screen monitor.

Outcome measures comprise the number of correct responses for each of the different delay types for easy and hard trials.

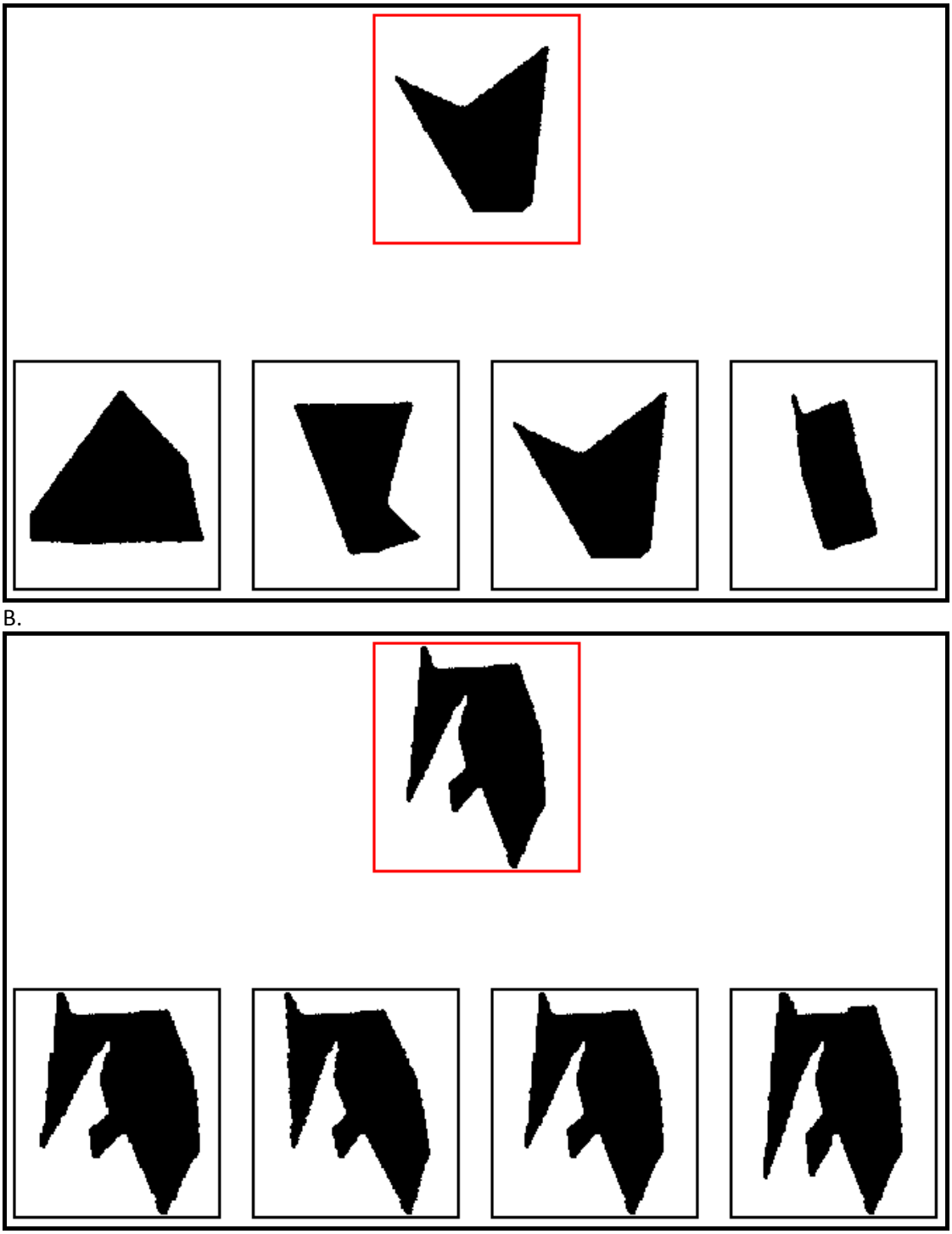


Figure 5.1: A) Example of an easy trial. Response options 1-3 are all 6-point shapes and option 4 is an 8-point shape.
B) Example of a hard trial with a 16-point shape. Option 3 is the correct answer. Option 1 is stretched horizontally by 15%; option 2 is skewed anti-clockwise by 8°; option 4 has had several parts of the image stretched.

12. Zoo Map (Behavioural Assessment of the Dysexecutive Syndrome (Wilson et al., 1996))

This subtest of the Behavioural Assessment of the Dysexecutive Syndrome presents the participant with a map of a pretend zoo, and they are instructed to draw a route around the zoo visiting six specific places out of the 11 that appear on the map. There are a number of restrictions on how the participant can travel around the zoo – for example the start and end point are fixed, and some paths may only be used once. In version A of the task, the participant is given the list of places to visit, along with the rules, and timed until they finish the route. In version B, the control task, the participant is given the order in which the places have to be visited, and again they are timed. For version A, there are a limited number of correct routes, and successful performance requires a degree of forward planning. Version B can serve as a psychomotor control task for those who successfully complete version A, or as a check that the participant is capable of drawing the route and following instructions for those who do not.

Both versions result in the same outcome measures: sequence score (the number of places visited in the correct sequence), planning time in seconds (time between the end of the instructions and the start of drawing the route), total time taken in seconds, errors (an amalgamation of the number of single-use paths used more than once, the number of inappropriate places visited, the number of failures to make a continuous line, and the number of deviations made), and a raw score (sequence score minus errors). The raw scores for versions A and B are added together and scaled to give an overall profile score which ranges from 0-4.

13. Digit Span (Corporation, 2002)

In the forwards digit span subtest (FDS), the participant is read a series of single-digit numbers at the rate of one per second and asked to repeat them in the same order

they were presented. The test starts with a series of three digits and two trials are administered at every level. The number of digits increases by one after each level where at least one of the trials is successfully recalled. The test continues until either two consecutive failures at the same level or until level 9 is reached, whichever comes first. Outcome measures comprise digit span (the longest correctly-recalled sequence), and score (the total number of trials correctly recalled).

For reverse digit span (RDS), as in FDS, the participant is read a series of digits at the rate of one per second, but is asked to repeat them in reverse order. The test starts with a series of two digits and proceeds as for FDS. Outcome measures comprise reverse digit span (the longest correctly-recalled sequence), and score (the total number of trials correctly recalled). Additionally a total score – the sum of the two scores for FDS and RDS – is calculated.

14. STROOP

This task has three parts. The first involves reading aloud colour names (red, blue, green, yellow) as quickly as possible to derive a reading time. The second involves saying aloud the colour of groups of 'X's (displayed in either red, blue, green, or yellow) as quickly as possible to derive a colour-naming time. The experimental task involves showing the participant the colour names written in incongruous colours (e.g. the word 'red' written in 'green'; red) and asking them to say aloud the colour the word is written in rather than the word itself.

Standard Stroop administration involves presenting the words in each condition on a sheet of card and timing the participant for the entire sheet. The present study used a computerised administration of the task. Words were presented individually on the screen and reaction times were logged by a voicekey responsive to sound intensity.

Whether the answers were right or wrong was recorded by the examiner. This administration gives a reaction time for each single word.

The task was run using Superlab 4.0 stimulus presentation software. The word-reading condition consisted of two throw-away trials followed by 16 trials of the words "yellow", "blue", "red", and "green" presented four times each in a random order (the throw away trials were not known to the participant and were not different in any way to the subsequent trials; the data for the first two trials of each condition were discarded to prevent average reaction time being skewed by initially slow reaction times which tend to occur in most participants at the beginning of this type of task). The words were displayed centred (both horizontally and vertically) in white font (Arial size 30) against a black background. For word-reading, as for all other trials, each participant saw the stimuli in the same random order. The trial was terminated as soon as the participant gave their response or after two seconds had elapsed, whichever was the sooner. Missed trials were not scored as errors, but were excluded from analyses of correct reaction times. There was a one-second intertrial interval before the next word appeared. At the beginning of the task participants were instructed to respond as quickly as possible (this instruction was not repeated at the beginning of each individual condition).

The colour-naming condition consisted of 32 experimental trials where a string of between three and six letter Xs were displayed in a given colour, either red, blue, yellow or green (e.g. 'XXX', 'XXXXX'). Each colour appeared twice at each of the four string lengths (3, 4, 5 or 6) to match the number of letters in each of the colour words. Trials were presented in a quasi-random order, subject to the constraint that each half of the 32 trials contained all possible trial types. The letter strings were displayed centred horizontally and vertically on a black background in capital letters (Arial size 30).

The colour-word condition consisted of 48 experimental trials where the words 'red', 'blue', 'yellow', and 'green' appeared written in incongruous colours. Each word appeared four times in each of the three non-matching colours. Once again, words were presented centrally in the relevant colour (Arial font size 30). Trials were quasi-randomised, such that each quarter of the task (12 trials) contained all of the possible trial types.

Participants wore a microphone mounted to a headset which detected their response. Responses were recorded as correct or incorrect by the person administering the test. For the purposes of analysis, average reaction times were recorded for each condition excluding extreme outliers (reaction time < 100ms or > 3 standard deviations above the participant's individual mean for that condition). The number of errors for each condition and the number of excluded values for each condition were recorded, and the interference scores (differences in average reaction time between word-reading and colour-naming, and between colour-naming and colour-word) were calculated.

PROCEDURE

The order of the sessions and the order of the measures are shown in full detail in Appendix 3 on page 312. For the neuropsychological testing, the battery was carried out in a single session for the majority of participants. For a small minority (n=5 (three controls and two patients)) testing was conducted over two sessions. The tests were administered in a fixed order for all participants (see Appendix 3) to maintain recall intervals for the delayed memory test and space the four digit symbol substitution tests as similarly as possible for all participants. The tests were divided into two blocks of approximately 70 minutes each with at least a 15 minute break between the two blocks. For those participants who were tested across two different sessions, the testing blocks were divided at the break such that the first session included the first 70 minute block of

tests and the second session – conducted as closely as possible to the first – contained the second block. For those participants who were tested over two sessions, the pattern of results on the repeat administrations of the DSST were compared with those who received the tests in a single session. The pattern was the same and therefore the results of all participants were included in the analysis of this task.

DATA ANALYSIS

Data were analysed using SPSS version 17.0. Demographics of the patient and control groups were characterised and compared using independent samples t-tests for continuous data or chi-squared tests for categorical data. The significance level was set to $p \leq 0.05$, with $p < 0.1$ representing a trend toward significance. Neuropsychological measures were compared using independent samples t-tests, or repeated measures ANOVA for tests that involved multiple levels or repetitions. For t-tests, Levene's F-test was first used to identify instances of unequal variance, and if $p < 0.05$ corrected p-values were reported. As part of data screening, the distributions of each variable were examined using boxplots. Any variables showing evidence of extreme outliers (values more than three times the inter-quartile range) were analysed using an appropriate nonparametric test (Mann-Whitney U test). If the results of the nonparametric test differed in terms of statistical significance from the parametric test, then the former was reported. The data were not formally tested for normality before using parametric techniques, as formal normality testing is less reliable in small samples. Additionally, many parametric techniques are not seriously affected by violations of assumptions (Glass et al., 1972).

To ascertain whether premorbid IQ or years of education may have been acting as suppressor variables and obscuring group differences on the neurocognitive measures, the neurocognitive outcome variables were correlated with premorbid IQ and

Table 5.3 : Demographic details and mood symptom scores for the patient and control groups

	Control			Patient					
	n	mean	s.d.	n	mean	s.d.	t/χ^2	p	d
Demographics									
Age	28	46.5	10.8	39	44.9	12.7	0.54	0.592	0.13
Male/female (%male/%female)									
Male (%)	13	(46.4)	-	18	(46.2)	-	0.00	0.982	-
Female (%)	15	(53.6)	-	21	(53.8)	-			-
Handedness									
Right (%)	26	(92.9)	-	34	(87.2)	-	0.61	0.737	-
Left (%)	1	(3.6)	-	3	(7.7)	-			-
Ambidextrous (%)	1	(3.6)	-	2	(5.1)	-			-
Nart IQ	28	114.4	8.9	37	110.4	10.7	1.62	0.110	0.41
Years of formal education	27	16.8	2.9	39	15.4	3.8	1.72	0.090	0.41
Smoker, n (%)	2	(7.1)	-	10	(25.6)	-	3.98	0.046	-
Cigarettes/day	2	6.0	5.7	9	13.4	9.1	-1.08	0.306	0.85
Alcohol units/week	28	7.9	9.5	38	7.4	8.3	0.22	0.830	-0.05
Screening Day									
Clinician-Rated Depressive Symptoms (HDRS-17)	28	0.9	1.2	38	3.8	2.1	-7.26	<0.001	1.68
Clinician-Rated Manic Symptoms (YMRS)	28	0.2	0.5	38	0.4	0.7	-1.26	0.212	0.30
Self-Rated Depressive Symptoms (BDI)	28	2.1	2.4	37	10.2	7.9	-5.88	<0.001	1.31
Self-Rated Manic Symptoms (AMRS)	28	1.7	2.6	38	2.3	2.9	-0.83	0.408	0.21
Self Esteem (Rosenberg)	28	24.9	4.1	38	16.4	5.9	6.53	<0.001	1.63
Dysfunctional Attitudes Scale									
Total Score	28	68.9	17.0	39	88.5	26.8	-3.65	<0.001	0.84
Achievement	28	21.0	8.4	39	29.1	10.6	-3.36	0.001	0.83
Dependency	28	22.5	5.6	39	30.3	9.4	-4.23	<0.001	0.97
Self-Control	28	25.6	7.3	39	31.2	11.9	-2.21	0.030	0.55
State Anxiety Inventory score	26	27.6	8.3	38	38.4	11.9	-4.02	<0.001	1.02
Trait Anxiety Inventory score	28	31.1	7.4	39	45.3	11.3	-6.20	<0.001	1.44
Test Day									
Clinician-Rated Depressive Symptoms (HDRS-17)	-	-	-	37	3.6	2.1	-	-	-
Clinician-Rated Manic Symptoms (YMRS)	-	-	-	38	0.6	1.5	-	-	-
Self-Rated Depressive Symptoms (BDI)	26	1.3	1.8	36	7.0	6.8	-4.76	<0.001	1.06
Self-Rated Manic Symptoms (AMRS)	28	2.1	3.3	36	1.8	3.2	0.44	0.665	-0.11

years of education using Pearson correlations. Group differences on outcome variables that were significantly correlated ($p<0.05$) with either of these demographic factors were compared again using ANCOVA with the relevant demographic factor included as a covariate. There were no instances when premorbid IQ or education was acting as suppressor variables (i.e. a previously non-statistically significant result reached statistical significance at $p<0.05$), therefore the results without the covariate are reported. The results of the ANCOVA analyses are reported in Appendix 4 on page 314. Effect sizes were calculated using Cohen's d where d=(control mean-patient mean)/pooled standard deviation. Effect sizes for indices where a higher score indicated worse performance (e.g. error or reaction time indices) were reflected (multiplied by -1) such that a positive effect size always indicates a poorer performance by the patient group. Conventions for Cohen's d are that $0.2<d<0.5$ is a small effect, $0.5\leq d<0.8$ is a medium effect, and $d\geq 0.8$ is a large effect. No explicit method was used to correct for multiple comparisons.

RESULTS

DEMOGRAPHICS

The demographic details of the patient and control samples are shown in Table 5.3 above. The groups were well-matched in that there were no significant differences between patients and controls in age, gender, education, or premorbid IQ. There were significantly more smokers in the patient sample than in the control sample.

With regard to symptoms, patients showed significantly higher levels of both clinician- and self-rated depressive symptoms at screening, and higher levels of self-rated depressive symptoms on the test day (n.b. controls did not have clinician-rated measures of symptoms on the test day, so it was not possible to compare levels of clinician-rated

depressive symptoms on test day). Despite this, the level of depressive symptoms was low in the patient group and within the range consistent with euthymia. There were no significant differences between patients and controls in clinician- or self-rated manic symptoms at any time point.

In the patient group, comparing symptom scores at screening to those four weeks later at test, there were no significant differences between clinician-rated depressive symptoms (t_{35}=0.41, p=0.68), clinician-rated manic symptoms (t_{36}=-1.0, p=0.32), or self-rated manic symptoms (t_{34}=0.65, p=0.52). There was a significant difference in self-rated depressive symptoms indicating that the patients had significantly lower scores (lower depression) on the test day compared to screening day (t_{34}=4.11, p=<0.001). However, this reduction was small (approximately 3 points on average). The interim self-rated mood scales confirmed this pattern (see Table 5.4), indicating patients remained relatively stable in the time between screening and test. There was no significant difference in self-rated manic symptoms (one-way repeated measures ANOVA, $F_{4,104}$=0.46, p=0.77), and a significant – but small – decrease in self-rated depressive symptoms ($F_{4,108}$=2.69, p=0.04). In all, there is no evidence of major mood change in the patient group over the time course of the study.

With regard to measures of dysfunctional attitudes, self-esteem, and state and trait anxiety, the patient group had significantly different scores to controls on every measure in the direction indicating higher psychopathology.

The details of clinical sample are provided in Table 5.4 below. For those in which it could be accurately ascertained, 19 (48.7%) had bipolar I disorder and 8 (20.5%) had bipolar II disorder. On average, patients had been diagnosed with bipolar disorder for approximately 15 years, although on average could track the onset of mood symptoms back to almost 8 years before diagnosis. Almost three-quarters of the sample (74.4%)

had been hospitalised at least once for their mood disorder. All of the patients were taking pharmacotherapy at the time of test.

Table 5.4 : Illness history, medication usage, and interim mood ratings of the Clinical Sample

	n	mean	s.d.
Diagnosis			
BDI, n (%)	19	(48.7)	-
BDII, n (%)	8	(20.5)	-
Unknown, n (%)	12	(30.8)	-
Illness History			
Age at onset of first mood episode	37	21.6	8.3
Age at diagnosis	36	29.3	10.8
Number of mood episodes in lifetime	35	17.2	20.2
Previously hospitalised for mood disorder, n (%)	29	(74.4)	-
Number of previous hospitalisations	29	4.9	3.8
Previously had ECT, n (%)	11	(28)	-
Time since last ECT (months)	11	195.8	91.2
Number of previous ECT treatments	10	7.3	11.6
Time stable on current medication (weeks)	31	74.8	83.7
Medication			
On Lithium	12	(30.8)	-
On an anti-depressant	19	(48.7)	-
On a typical antipsychotic	3	(7.7)	-
On an atypical antipsychotic	19	(48.7)	-
On a benzodiazepine	3	(7.7)	-
Has received psychological therapy	13	(33.3)	-
Interim Mood Ratings			
Self-Rated Depressive Symptoms - Week 1 (BDI)	34	9.4	7.1
Self-Rated Depressive Symptoms - Week 2 (BDI)	32	8.5	6.6
Self-Rated Depressive Symptoms - Week 3 (BDI)	33	8.6	7.9
Self-Rated Manic Symptoms - Week 1 (AMRS)	34	1.8	2.3
Self-Rated Manic Symptoms - Week 2 (AMRS)	32	2.1	2.7
Self-Rated Manic Symptoms - Week 3 (AMRS)	32	2.4	3.5

NEUROPSYCHOLOGICAL FUNCTION

Results of the neuropsychological tests are reported in Table 5.5 through to Table 5.12. The verbal memory measures are reported in Table 5.5. Patients showed significantly impaired performance on some indices of the Rey Auditory Verbal Learning Test (RAVLT). Specifically, patients showed impaired recall on the first presentation of list A and impaired total recall. There was a trend towards a significant difference in short-delay recall. There were no significant differences in long-delay recall, forgetting, or recognition.

The effect sizes reported here for total recall (d=0.5), short-delay recall (d=0.46), and long-delay recall (d=0.17) are smaller than those reported in the meta-analyses (see Chapter 2, Table 1.1, on page 32). This may relate to the different format of the task as executed in the present study. Comparing the patient group data to that in a previous

Table 5.5: Results of verbal memory measures

MEMORY - VERBAL	Control		Bipolar		t	df	p	d
	mean	s.d.	mean	s.d.				
Rey Auditory Verbal Learning Test (RAVLT)								
Total recall trials 1-5	**47.30**	**6.54**	**43.58**	**7.94**	**2.00**	**63**	**0.050**	**0.50**
A1	**6.39**	**1.64**	**5.50**	**1.56**	**2.25**	**64**	**0.028**	**0.56**
B	5.54	2.10	5.18	1.67	0.76	64	0.452	0.19
A6 (Short-delay)	9.36	2.42	8.26	2.77	1.67	64	0.099	0.42
A7 (Long-delay)	8.57	2.50	8.08	3.17	0.68	64	0.499	0.17
Forgetting (% retained from A6 on trial A7)	94.76	18.67	97.50	26.89	-0.46	64	0.644	-0.12
Recognition list A hits	12.86	1.67	12.32	2.41	1.02	64	0.311	0.25
Wechsler Memory Scale Logical Passages Test								
A+B recall units (0-50)	**28.18**	**7.55**	**23.05**	**6.86**	**2.86**	**63**	**0.006**	**0.72**
A+B thematic units (0-14)	**11.64**	**2.20**	**10.41**	**2.54**	**2.06**	**63**	**0.044**	**0.52**
Forward Digit Span								
Span	7.18	0.98	6.95	1.21	0.83	64	0.409	0.21
Score	8.93	2.02	8.63	2.31	0.54	64	0.588	0.14

comparable study using standard administration (Thompson et al., 2005) and the control group performance to normative data (Strauss et al., 2006), it seems that presenting the word list in a different order on each trial has impacted on performance in both groups (Thompson et al 2005, A1-A5 patient mean=46.7, s.d.=9.1, 107% higher than present sample; Strauss et al 2006 A1-A5 normative data for age 40-49 mean=51.1, s.d.=8.6, 108% higher than present sample). However, in relative terms, the control group has been more affected (see Chapter 6 for further discussion of this issue).

In the story recall task patients demonstrated impairment on both recall of story units and recall of story themes. The effect size (d=0.72) is higher than the most recent meta-analysis, in which this index was meta-analysed for the first time (d=0.63) (Kurtz and Gerraty, 2009), but both are medium effect sizes.

There was no evidence of impairment in immediate memory as assessed with forwards digit span, with a small effect size of 0.21. The meta-analyses have also reported small effect sizes in the region of 0.4 for forwards digit span, and, in line with the present findings, it tends to be one of the lower effect sizes reported.

Table 5.6: Results of visual memory measures

MEMORY - VISUAL	Control		Patient		t	df	p	d
	mean	s.d.	mean	s.d.				
Simultaneous & Delayed Match to Sample								
"Easy" trials								
Total Correct	**16.93**	**0.81**	**16.27**	**1.37**	**2.42**	**60**	**0.019**	**0.57**
Correct (simultaneous trials)	6.00	0.00	6.00	0.00	-	-	-	-
Correct (0-sec delay trials)	**5.89**	**0.32**	**5.54**	**0.65**	**2.88**	**55**	**0.006**	**0.66**
Correct (4-sec delay trials)	5.04	0.74	4.73	1.02	1.40	63	0.167	0.34
"Hard" trials								
Total Correct	11.64	2.26	11.08	2.45	0.94	63	0.348	0.24
Correct (simultaneous trials)	4.11	1.10	4.24	1.21	-0.47	63	0.643	-0.11
Correct (0-sec delay trials)	3.75	1.14	3.51	1.07	0.86	63	0.395	0.22
Correct (4-sec delay trials)	3.79	1.32	3.32	1.44	1.33	63	0.188	0.34

Results of the visual memory task are presented in Table 5.6. Although there were no differences between groups on the hard trials of the task, patients showed significantly worse performance than controls on the 0-second delay easy trials. Given that the patients were not impaired in the 4-second delay trials, or on the hard trials, this seems unlikely to be a genuine failure of visual memory. It is possible this was due to guessing or impulsive responding.

As reported in Table 5.7 below, patients showed impairment in verbal fluency, but only in the generation of words by category. The effect size (0.72) is consistent with the most recent meta-analysis (0.75) (Kurtz and Gerraty, 2009). Although patients were worse than controls for phonetic fluency, this difference did not reach statistical significance and the small effect size reported (0.39) is consistent with earlier meta-analyses. There was a trend towards the patients making significantly more rule breaks than controls during the phonetic fluency task.

Table 5.7: Results of verbal fluency measures

EXECUTIVE - FLUENCY	**Control**		**Patient**		**t**	**df**	**p**	**d**
	mean	**s.d.**	**mean**	**s.d.**				
Verbal Fluency								
FAS								
Total correct	47.96	11.96	43.29	11.97	1.57	64	0.122	0.39
Total rule breaks	0.43	0.69	0.86	1.11	-1.95	61	0.056	0.46
Total repeats	1.32	1.39	1.08	1.26	0.74	64	0.462	0.18
Category								
Total correct	**67.68**	**13.01**	**56.92**	**16.32**	**2.89**	**65**	**0.005**	**0.72**
Total repeats	1.36	1.37	1.10	1.68	0.66	65	0.512	-0.16

Results for the tasks of inhibition are presented in Table 5.8 below. There were trends towards the patients making more category B errors and responding more slowly than controls on part 2 of the Hayling Sentence Completion Test. The effect sizes were small (around 0.45). This test has yet to be subject to meta-analysis in this population, but this effect size is similar to that found by another study (d=0.52) (Dixon et al., 2004).

Table 5.8: Results of executive tests measuring inhibition

EXECUTIVE - INHIBITION	Control		Patient		t	df	p	d
	mean	s.d.	mean	s.d.				
Hayling Sentence Completion Test (HSCT)								
Total Time part 1 (scaled)	5.75	0.75	5.77	0.71	-0.11	65	0.915	-0.03
Total Time part 2 (scaled)	6.14	0.45	5.90	0.64	1.74	65	0.086	0.42
A score	2.79	4.38	3.00	5.31	-0.17	65	0.862	0.04
B Score	3.86	5.52	8.13	11.53	410.5[a]	58	0.082	0.45
Overall Scaled Score	5.96	1.17	5.54	1.48	1.26	65	0.212	0.31
Stroop								
Reading time (average ms/word)								
Word Reading	497.07	77.27	509.92	82.80	-0.62	59	0.538	0.16
Colour Naming	627.59	84.86	648.62	101.08	-0.87	59	0.391	0.22
Colour-Word	834.31	158.60	887.07	210.28	-1.08	59	0.284	0.28
Errors								
Word Reading	0.00	0.00	0.00	0.00	-	-	-	-
Colour Naming	0.15	0.46	0.26	0.51	-0.93	59	0.357	0.24
Colour-Word	**0.19**	**0.48**	**1.15**	**1.50**	**-3.52**	**41**	**0.001**	**0.82**
Total	**0.67**	**0.73**	**2.40**	**2.58**	**-3.78**	**41**	**0.001**	**0.87**
Excluded values								
Word Reading	0.37	0.63	0.32	0.73	0.27	59	0.792	-0.07
Colour Naming	0.85	1.20	0.47	0.79	348.0[a]	59	0.069	-0.38
Colour-Word	0.96	0.81	1.47	1.33	-1.84	56	0.071	0.45
Total	3.15	2.18	3.36	2.42	-0.36	58	0.721	0.09
Colour word minus word reading	337.23	167.81	377.15	194.46	-0.85	58	0.401	0.22
Colour word minus colour-naming	206.71	109.44	238.45	145.04	-0.94	58	0.349	0.24

[a] Mann-Whitney U

In the Stroop test, patients made significantly more errors in the colour-word trial than controls. They also showed a trend towards having significantly more excluded reaction times in the colour-word condition (although it should be noted that there was a trend for the controls to make more errors in the colour naming condition). Values were excluded for very short or relatively long response times. This was designed to minimise the influence of outliers, but may reflect greater difficulty with the task for the bipolar patients. In the standard administration of the Stroop, time taken and errors can be confounded (if someone hesitates then makes an error, or makes an error which they then correct, the extra time is reflected in the time taken for the task as a whole). In the

present task, as errors are excluded from the reaction time data, it becomes clear that patients are not significantly slower at the task, but they do tend to make more mistakes. The effect reported here (d=0.82) is similar to, if not slightly larger than, that reported in the meta-analyses (d=0.75, Kurtz and Gerraty, 2009), perhaps reflecting the better separation of errors and response time.

Table 5.9: Results of executive tests of planning

EXECUTIVE - PLANNING	Control		Patient		t	df	p	d
	mean	s.d.	mean	s.d.				
BADS Zoo Map								
Version 1 sequence score	5.93	2.42	4.60	2.98	1.88	60	0.065	0.48
Version 1 planning time (seconds)	93.70	93.89	80.17	115.84	0.49	60	0.623	-0.13
Version 1 total time (seconds)	169.89	112.77	174.23	121.50	-0.14	60	0.886	0.04
Version 1 total errors	1.11	1.93	3.06	4.73	366.5[a]	47	0.106	0.51
Version 1 raw score	**4.81**	**3.75**	**1.69**	**6.81**	**2.30**	**55**	**0.025**	**0.55**
Version 2 sequence score	7.36	1.97	7.77	1.06	-1.00	39	0.321	-0.27
Version 2 planning time (seconds)	15.82	25.96	10.43	16.51	1.00	61	0.320	-0.25
Version 2 total time (seconds)	52.82	31.45	58.06	31.47	-0.66	61	0.514	0.17
Version 2 total errors	0.04	0.19	0.71	2.83	-1.41	34	0.167	0.32
Version 2 raw score	7.32	2.07	7.29	3.57	0.05	61	0.963	0.01
Overall profile score	2.81	1.47	2.15	1.54	1.72	59	0.091	0.44
Self-Ordered Pointing Test (SOPT)								
Total errors all levels	**9.14**	**4.44**	**12.69**	**6.50**	**-2.50**	**65**	**0.015**	**0.62**
Total errors level 4	0.75	0.93	0.87	0.89	-0.54	65	0.590	0.13
Total errors level 6	**1.54**	**1.64**	**2.44**	**1.71**	**-2.16**	**65**	**0.035**	**0.53**
Total errors level 8	**2.93**	**1.54**	**4.00**	**2.49**	**-2.17**	**64**	**0.034**	**0.50**
Total errors level 10	**3.93**	**2.37**	**5.38**	**2.70**	**-2.29**	**65**	**0.025**	**0.57**
Highest level with at least 1 trial correct	7.93	2.21	7.08	2.55	1.42	65	0.159	0.35
Maximum span level 4	3.96	0.19	3.97	0.16	-0.24	65	0.815	-0.06
Maximum span level 6	5.64	0.56	5.41	0.94	1.27	63	0.210	0.29
Maximum span level 8	**7.32**	**0.90**	**6.69**	**1.38**	**2.25**	**65**	**0.028**	**0.52**
Maximum span level 10	**8.89**	**1.20**	**8.00**	**2.12**	**2.19**	**62**	**0.032**	**0.50**

[a] Mann-Whitney U

Table 5.9 above reports the results of the planning tasks. Patients showed impairment in the Zoo Map task, performing significantly worse overall than controls on the first part of the task. Patients showed no impairment on part two, the control task, suggesting there were not basic difficulties in route-drawing or following instructions

that could explain the differences. In both versions of the map task patients took less time to plan, however this difference was not statistically significant.

Patients showed impaired performance on the self-ordered pointing task, making more errors than controls at every level of the task apart from the first. All in all, the effect sizes reported were moderate. This task is grouped under the heading of planning tasks, as successful performance involves a degree of forward planning to determine in which order to touch the shapes. However, the task is incredibly complex in terms of the processing it involves (for example it also involves the continuous monitoring and updating of working memory). It is therefore difficult to conclude on the basis of elevated errors that patients have an impairment specifically in planning. Table 5.9 also reports span scores – the maximum number of correct patterns selected before an error was made – which are significantly reduced in the patient group for levels 8 and 10. Patients may have failed to use an optimal strategy, or they may have reached their capacity limit sooner than controls.

Table 5.10: Results of measures of mental manipulation

EXECUTIVE - MENTAL MANIPULATION	Control		Patient		t	df	p	d
	mean	s.d.	mean	s.d.				
Reverse Digit Span								
Span	5.50	1.20	5.08	1.38	1.29	64	0.201	0.32
Score	7.50	2.06	7.00	2.27	0.92	64	0.361	0.23
Total Score (forwards score + reverse score)	16.43	3.76	15.66	3.84	0.81	64	0.419	0.20

Table 5.10 above reports the results of the mental manipulation measure. There is no evidence of impairment in reverse digit span. The small effect size reported here (0.32) is lower than has been reported in the meta-analyses to date, which have shown moderate-to-high effect sizes on this measure ($0.54<d<1.09$). However, it is one of the indices which has shown a great deal of variation between studies, and the result

reported here is not out of line with some other studies in euthymic patients (Kieseppä et al., 2005, Thompson et al., 2005).

Table 5.11: Results of set-shifting measures

EXECUTIVE - SET-SHIFTING	Control		Patient		t	df	p	d
	mean	s.d.	mean	s.d.				
Trail Making Test								
TMTb time (sec)	68.38	46.44	76.55	36.68	-0.80	64	0.427	0.20
Switch (TMTb-TMTa)	40.35	46.03	40.84	31.27	-0.05	64	0.959	0.01

There was no evidence of set-shifting impairment as measured by part B of the trail making test (Table 5.11 above). The meta-analyses have reported moderate to large effect sizes between $0.55<d<0.86$ for this measure, much larger than the small effect reported here ($d=0.2$). Comparing the present result to the most comparable similar study (Thompson et al., 2005), the mean difference between the groups for total time is very similar (8.17 seconds here versus 5.8 seconds in Thompson et al), however there is much more variation in both the patient and control groups in the current study (approximately 20 seconds in the control group and 10 in the patient group), possibly due to the smaller sample size, which has reduced the effect size.

Table 5.12: Results of psychomotor and learning capacity measures

PSYCHOMOTOR & LEARNING CAPACITY	Control		Patient		t	df	p	d
	mean	s.d.	mean	s.d.				
Digit Symbol Substitution Test (DSST)								
Attempt 1	54.36	11.53	48.26	11.87	2.10	65	0.040	0.52
Attempt 2	62.43	11.83	51.64	13.15	3.45	65	0.001	0.85
Attempt 3	64.75	12.19	54.70	12.56	3.23	63	0.002	0.81
Attempt 4	67.00	13.20	57.14	13.15	2.97	62	0.004	0.75
% change from 1-2	15.71	10.80	7.31	11.17	3.08	65	0.003	0.76
% change from 2-3	4.09	7.90	6.26	9.11	-1.01	63	0.318	-0.25
% change from 3-4	3.62	7.44	3.43	7.91	0.10	62	0.920	0.03
% change from 1-4	24.48	14.79	16.85	15.52	1.99	62	0.051	0.50
Learning slope	4.03	2.40	2.56	2.30	2.48	62	0.016	0.61
Trail Making Test								
TMTa time (sec)	28.02	9.36	35.71	13.93	-2.70	65	0.009	0.63

The results of the psychomotor measures are reported in Table 5.12 above. Patients showed impairment relative to controls on each administration of the digit symbol substitution test. However, the relative change in performance with each subsequent administration only showed a difference between patients and controls for the percentage change between the first and second administrations. This suggests that although patients start from a lower baseline, and initial learning is impaired, subsequently they gain at approximately the same rate as healthy controls. The medium effect size reported for the first administration of the digit symbol substitution test (d=0.52) is marginally lower but broadly consistent with those reported in the meta-analyses ($0.59 \leq d \leq 0.84$). Patients showed a significantly flatter learning slope than controls, indicating that on average across the repeat administrations they made performance gains at a slower rate.

Patients showed impairment in part A of the trail making test. The medium effect size reported here (0.63) is consistent with the meta-analyses (range 0.52-0.69).

SUMMARY OF RESULTS

In summary, patients showed no evidence of impairment in forwards digit span (a measure of immediate verbal memory), generational fluency by letter, mental manipulation, or set-shifting as measured by the trail making test. On all other measures patients demonstrated significantly worse performance than controls on at least one test index (verbal memory, visual memory, generational fluency by category, inhibition, planning, psychomotor speed, and learning capacity).

DISCUSSION

Consistent with a large number of previous studies, patients with bipolar disorder showed impaired performance on a number of neuropsychological tests. In terms of broad categories, this comprised impairment on tests of verbal memory & learning, visual memory, inhibition, planning, verbal fluency, psychomotor processing, and learning capacity. The only domains in which no statistically significant impairments were evident were set-shifting (as measured by the trail making test – see Chapter 7 for further details of a different set-shifting task) and mental manipulation (although note that patient performance was nonetheless below that of controls on both of these measures).

In terms of the size of the effects reported here, the pattern is broadly in line with the effect sizes reported in the meta-analyses. The largest effect sizes reported here are for aspects of executive function – inhibition (from the stroop task; d=0.82) and fluency (by category; d=0.72) – which matches the results of the most recent meta-analysis very closely (d=0.75 and d=0.75 respectivley; Kurtz and Gerraty, 2009). In terms of the lowest effect sizes, in the present sample delayed verbal recall showed the smallest effect size (d=0.17), which is in contrast with the meta-analysis where, of the overlapping tests, forwards digit span shows the smallest effect size (d=0.41; effect size for delayed verbal recall d=0.78). As mentioned above, differences in the administration procedure for the verbal learning test may account for the lower effect sizes in this task. Specifically, the present study used a standard list-learning paradigm, but shuffled the presentation order of the list for each of the five presentations. This altered the demands of the task (as intended), but comparing the mean scores to normative data and previous studies shows that the performance of the control sample was relatively more affected than the performance of patients. The lower effect sizes seem to reflect

poorer performance by the control group rather than better than average performance by the patient sample (see Chapter 6 for further details). Bearing this in mind, if the effect sizes for the verbal memory task are ignored for the purposes of comparison with the meta-analyses, then the lowest effect size reported presently is forwards digit span (d=0.21). Although lower than the most recent meta-analysis, both are small effects and both rank in the same order of all comparable tests administered.

It should be noted briefly that one of the largest effect sizes here – category fluency – is described as an executive deficit. However, there is debate about whether category fluency is an executive function. Evidence indicates that it taps into temporal semantic networks and is sensitive to deficits in semantic memory (Mummery et al., 1996, Rosser and Hodges, 1994). The impairment observed in patients with bipolar disorder on this measure may reflect a disturbance in semantic memory processes rather than executive processes per se.

Two of the effects were considerably lower than reported in the meta-analyses – set-shifting as measured with the trail making test (d=0.20 in contrast with d=0.73), and mental manipulation as measured by reverse digit span (d=0.32 in contrast with d=0.65). However, neither result was out of keeping with previous individual studies and reflects relatively large variation even between meta-analyses (which essentially all draw from the same core pool of studies with only a few extra studies in each subsequent review) for these indices. It may be that these indices are sensitive to variation from other sources and thus the differences between studies reflect not only the effects of bipolar disorder, but also other key differences between samples (e.g. age, illness history, educational level etc.). One important point to note in the present sample is that inter-rater reliability on the diagnostic and symptom measures was not established. It was therefore not possible to identify whether any potential variation in the clinical sample

resulting from slight differences in assessment practices impacted on the present results. It is possible that some of the differences noted between the current sample and previous studies relate to differences in how patients were assessed. However, as the diagnostic instruments were used to confirm an already-existing diagnosis rather than to establish a new diagnosis, it is expected that any impact on the results is minor.

The above comparison establishes that the sample investigated in the current study is comparable with previous similar studies.

CHAPTER 6: MEMORY AND EXECUTIVE FUNCTION IN BIPOLAR DISORDER

INTRODUCTION

The aim of the current chapter is to explore in more depth the relationship between executive function impairment and verbal memory impairment in bipolar disorder. The question of whether the many and varied differences noted between patients and controls stem from a single difficulty in a cognitive process common to all of the impaired tasks, or whether they represent multiple different underlying deficits is an issue that is unresolved in the literature.

With regard to understanding the underlying nature of cognitive dysfunction in patients with bipolar disorder, only a very small number of studies have made an attempt to investigate this issue. Many of the studies of cognitive function in this patient group have focused on phenomenological questions, which – reduced to their fundamentals – simply ask whether deficits are present at all and perhaps whether particular subsets of patients (or their relatives) demonstrate a broader or narrower profile of impairments. Speaking generally, the primary aim has been to establish what part cognitive function plays in the overall mystery of bipolar disorder, with particular focus on whether assessment of cognitive deficits could play a role in assisting diagnosis/identification of the illness, identifying subgroups for genetics studies, and whether it can be used as an outcome measure in treatment studies.

Viewpoints vary as to how successful measuring cognitive function will be at any of these roles (for opposing views see (Glahn et al., 2004) and (Savitz et al., 2005)), but

one limiting factor is a lack of clarity as to the actual nature of the impairment(s). The interpretation of the results of neuropsychological investigations in bipolar disorder has generally not accounted for the 'dirty' nature of neuropsychological tests. Although many tests have a well-established history and strong empirical foundation, all neuropsychological tests involve multiple functions. Ascertaining which process is impaired and underlies a poor performance requires detailed investigation. When patient groups show evidence of impairment across a range of tasks, failure of one process or one piece of neural hardware may be responsible, or several different underlying impairments may exist. To differentiate these alternatives, it is necessary to investigate beyond the existence of between-group differences, and begin to consider the relationship between performance on different types of tasks.

This is a large undertaking that requires an extended series of studies. It will not be solved within a single investigation, and the current chapter examines only one small piece of the broader puzzle. A small number of authors have attempted to understand more about the processes underlying performance on specific tasks (Deckersbach et al., 2004a, Deckersbach et al., 2004b, Thompson et al., 2007) or about patterns of performance on a small group of tasks using a model-driven approach (Thompson et al., 2006). In a pair of similar studies, Deckersbach and colleagues used path analysis to investigate whether strategy-use accounted for recall deficits on a verbal list-learning task (the California Verbal Learning Test (CVLT)) (Deckersbach et al., 2004b) and a visual figure recall task (Rey-Osterrieth Complex Figure (ROCF)) (Deckersbach et al., 2004a). Both of these tasks provide scoring indices reflecting the use of strategy. For the CVLT a 'semantic clustering' index can be calculated which reflects the extent to which an individual's recall is organised into the four semantic categories that comprise the CVLT word list. For the ROCF a 'copy organisation' score can be derived, which reflects the extent to which the major structures of the complex figure were used during recall.

Although both these tasks measure the ability to remember information, memory is not only limited by the capacity of the store available. The efficiency with which information is entered into that store and then retrieved from it also has some bearing on how much is recalled (Gathercole, 1999). In both of the studies, the patient group showed significantly lower immediate recall than controls. The authors went on to investigate whether the relationship between group (i.e. patient or control) and recall was mediated by the degree of semantic clustering/copy organisation using path analysis. For the verbal recall task, although semantic clustering was a significant mediator of performance, it did not explain all of the variance in recall (Deckersbach et al., 2004b). In contrast, for the visual recall task, impairment in immediate recall was fully mediated by copy organisation and the direct relationship between group and recall was statistically non-significant in the mediated model (Deckersbach et al., 2004a). Both of these studies show that an executive component of memory – the use of structure or organisation – shows impairment in patients with bipolar disorder. Further to that, the impairment in this executive component has a significantly negative impact on performance. Deficits on the two memory tasks were a reflection of executive difficulties rather than solely a reflection of reduced memory capacity or impaired memory processes per se.

Investigating a different aspect of cognitive function, spatial working memory, Thompson and colleagues used an approach guided by Baddeley's working memory model to identify factors contributing to impaired performance (Thompson et al., 2006). The computer analogue of the Corsi Block Test, the CANTAB Spatial Span test, has shown impairment in patients with bipolar disorder when euthymic (Thompson et al., 2005). Using analysis of covariance, Thompson and colleagues demonstrated that differences between euthymic patients and controls on the spatial task were no longer statistically significant once executive function was included as a covariate. The same was true when different executive measures were used as covariates, including an executive index

derived using factor analysis. The authors concluded that impairment on the spatial task was more likely to be executive in nature rather than a failure of short term spatial memory.

In a separate study on data derived from an overlapping sample, the same group analysed performance on a visual self-ordered pointing task to identify possible explanations for the impairment shown on this task (Thompson et al., 2007). Four hypotheses were derived, and a number of additional performance indices for this task were calculated to identify which hypothesis showed the best fit. The results showed that the deficit in self-ordered pointing task performance was best explained by a deficit in the monitoring and updating of working memory. However, in this study, two of the remaining three hypotheses tested had statistical support at trend levels ($p<0.1$). It is necessary to place a cautious interpretation on these results until the findings are replicated (it is possible that the hypotheses were not truly independent and/or that the performance indices were inter-related). However, this study demonstrates the principle that further exploration of initial findings is fruitful for drilling down to develop a better understanding of the nature of underlying processing difficulties.

The focus of the present study is on the relationship between verbal memory and executive functioning. As was established in Chapter 2, meta-analyses have reported that verbal memory and aspects of executive function show the greatest degree of impairment in patients with bipolar disorder. Given the strategic and organisational aspects of verbal learning, it may be that these two areas are related to one another. One study in a mixed group of patients referred for neuropsychological assessment reported that executive function and memory shared 55-60% of variance in a canonical correlation analysis (Duff et al., 2005).

Data in patients with major depression has indicated a relationship between verbal memory impairment and executive function (Fossati et al., 2002). There is also a positive relationship between the number of previous episodes of depression and the degree of memory impairment (Fossati et al., 2004, Gorwood et al., 2008), suggesting that memory processes may be especially sensitive to the effects of depression. In patients who are currently depressed, the factor that plays the greatest explanatory role in the degree of memory impairment demonstrated is the severity of depressive symptoms (Gorwood et al., 2008). Once patients are in remission, illness history factors predominate (Gorwood et al., 2008). In patients with bipolar disorder, the clinical data also indicate a greater burden of illness is associated with greater cognitive dysfunction. One review of studies that correlated neuropsychological dysfunction with illness history indices in patients with bipolar disorder reported relatively more consistent evidence for a relationship between memory measures and various features of illness including length of illness, number of manic episodes and number of hospitalizations (although note the relationship in patients with bipolar disorder is between manic episodes and memory function rather than depressive episodes and memory function (Robinson and Ferrier, 2006)). On the other hand, executive function did not show a consistent relationship with illness features (Robinson and Ferrier, 2006). First degree relatives of individuals with bipolar disorder show evidence of executive impairment, but very limited evidence of memory impairment (Bora et al., 2009). Likewise, there is evidence of executive impairment very early in the course of bipolar disorder, but the one study that has contrasted the cognitive function of first versus multi-episode patients did not report deficits in memory function (Nehra et al., 2006).

The evidence is difficult to reconcile. On the one hand there is evidence that executive aspects of verbal and visual memory processing show deficits in patients with bipolar disorder and that these deficits impact negatively on recall. On the other hand

there is evidence that executive function is impaired early in the illness without notable memory impairment, but impaired memory develops over time with repeated episodes. It could be that patients are able to compensate for impaired executive function early on in the course of their illness, but over time the deleterious effects of episodes limits the use of these compensatory strategies and exposes the core executive deficit. This would lead toward the working hypothesis of the present study that the verbal memory impairment in euthymic patients with bipolar disorder is primarily driven by executive function impairment. The strategic aspects of memory which support effective encoding or retrieval are impaired, which is sufficient to explain the deficit noted in euthymic patients. Whereas in depressed patients additional impairment will be caused by the effects of depression and controlling for any differences in the executive aspects of memory will not be sufficient to explain the full extent of the deficit they show.

The verbal memory paradigm most frequently investigated in bipolar disorder has been free recall of a supraspan word list. Presenting the same list of words multiple times to the same participant allows assessment of both their capacity for retaining recently-presented material, and also the degree to which they can learn and recall more items with each subsequent presentation-recall opportunity. The short-term store for phonological information suffers a significant capacity constraint (as a rule of thumb 7 items plus or minus two (Miller, 1956b)), the proposed organisation and chunking of information that takes place over the course of learning (Miller, 1956a, Miller, 1956b) adds an additional source of variation in performance. This process, a reasonable candidate for an executive component of verbal learning and memory, may be reflected in a change in the degree of organisation of a participant's recall output with each subsequent recall. Two phenomena of particular interest that have been noted in multi-trial free recall tasks are 1) participants recall a greater number of items on each trial, and 2) recall output becomes more organised (Sternberg and Tulving, 1977), that is the

same items are recalled in adjacent positions with increasing frequency across trials. In appropriate test formats (i.e. those in which the word list is presented in a different order for each trial), this emerging organisation is termed subjective organisation, since it reflects organisation imposed by the learner (generally, though not necessarily, on the basis of the semantic or phonological properties of the list items) rather than due to the presentation order of the list.

As a concept itself, subjective organisation has received much attention throughout the history of verbal learning research. However, it has yet to be investigated in patients with bipolar disorder. The key premise rests on the assumption that, under free-recall conditions, the items recalled by participants are not a random selection of all items presented and properties of the recall (especially comparisons of the properties of the recalls across different trials of the same task by the same participant) shed some light on the processing of inputted material that takes place (Pellegrino and Hubert, 1982). Subjective organisation correlates positively with overall recall (Pellegrino and Battig, 1974, Tulving, 1964), potentially indicating that this organisational process is an aide to recall.

There are a number of issues that surround quantifying subjective organisation and there are several different indices that can be used (Bousfield and Bousfield, 1966, Pellegrino, 1971, Tulving, 1962). The basis of most measures is that increased organisation is reflected in a greater number of pairwise adjacencies reoccurring with each additional recall (i.e. two words recalled next to one another in the recall output, also termed an intertrial repetition). The indices are generally calculated using a pair of consecutive recall trials of the same list by the same participant. The two most commonly used measures are Pellegrino & Pellegrino's ARC' (Adjusted Ratio of Clustering (Pellegrino, 1971)) and Tulving's SO (Subjective Organisation (Tulving, 1962)).

These two measures suffer less from some of the pitfalls inherent in many other measures. Two of the key issues are mathematical independence between the subjective organisation measure and total items recalled, and adjustment for chance-level performance. With regard to the first issue, many of the available measures are limited by an artefactual relationship between the subjective organisation index and the number of items recalled (Murphy and Puff, 1982). This tends to produce the situation where lower levels of recall are associated with less organisation whether or not this relationship truly exists in the data. Conceptually, this difficulty stems from the fact that increasing the number of items recalled in two separate trials also increases the number of items recalled in common between the two trials. In turn, the number of pairwise adjacencies likely to occur by chance increases. The ultimate result is that failure to account for these artefacts can produce an artificially high relationship between subjective organisation and recall, as well as a mis-estimation of the true level of organisation. With regard to the second issue – accounting for chance level of organisation – failure to adjust for chance expectancies results in a subjective organisation measure which is unbounded (i.e. can take any value up to the length of the list itself) therefore the magnitude of the index becomes difficult to understand. Comparing between studies with different list-lengths is problematic, and comparing between subjects who recalled different numbers of words is likewise problematic. It also becomes especially difficult to interpret negative values. One of the key features of a subjective organisation measure is that it should be possible to interpret the degree of organisation relative to that which would have been expected by chance. The two measures ARC' and SO each adjust for at least one of these issues.

Pellegrino & Pellegrino's ARC' measure is calculated as in equation (1). The denominator calculates the difference between the observed number of repeated pairwise adjacencies (intertrial repetitions) and the number that would be expected by

chance (expected number of intertrial repetitions; calculated assuming random ordering of the items that were recalled (i.e. it does not account for the fact that the items recalled were a subset of the total number of items presented)). The denominator calculates the difference between the maximum possible number of intertrial repetitions and the expected number. This index is bounded between -1 and +1, with 0 representing chance organisation. The closer the individual's organisation is to the maximum possible, the closer the numerator and denominator will be, and therefore ARC' score will be closer to +1. In other words, positive values indicate higher organisation, with +1 representing perfect organisation (the overlapping items of the two lists were recalled in exactly the same order). Negative values represent below-chance levels of organisation.

$$ARC' = (OITR-EITR)/(MAXITR-EITR) \quad (1)$$

Where:

OITR = Observed intertrial repetitions

EITR = Expected intertrial repetitions

MAXITR = Maximum intertrial repetitions

The more straightforward calculation of Sternberg & Tulving's SO is shown in equation (2). The observed number of intertrial repetitions is corrected for the number of common items recalled between the two list recalls. By correcting for the number of common items, this measure automatically adjusts for differences in the number of items recalled between individuals. The index itself can take values between 0 and (just less than) 1, with 0 representing no organisation (or no common items recalled), and higher values representing greater organisation. This measure does not correct for chance-level performance.

$$SO = OITR/(C+1) \quad (2)$$

Where:

OITR = Observed intertrial repetitions

C = Number of common items recalled on trials *t* and *t+1*

Studies investigating the degree of subjective organisation in patients with brain damage have shown that patients with frontal lobe injury show lower levels of subjective organisation than healthy controls (Gershberg and Shimamura, 1995, Heubrok, 1999, Stuss et al., 1994). Stuss et al (1994) reported that – irrespective of the laterality or extent of the lesion – patients with frontal lobe lesions showed lower levels of subjective organisation than healthy controls. Within groups, the only difference noted was that patients with a unilateral right-sided lesion showed a greater degree of subjective organisation for a word list with words drawn from a discrete number of categories (blocked) as contrasted with an unstructured list. Patients with a unilateral left-sided lesion and patients with bilateral lesions showed no evidence of using the categories on the blocked list to aid their organisation.

With respect to bipolar disorder, subjective organisation has yet to be investigated in this patient population. Additionally, the list-learning tests most often used to assess memory have been administered in the standard way – with fixed presentation of the word list across five consecutive trials. The present study investigates first whether there is any evidence of differences in subjective organisation in depressed patients with bipolar disorder or euthymic patients with bipolar disorder using the standard administration of a list-learning task. It then goes on to investigate subjective organisation in euthymic patients with bipolar disorder using a shuffled list presented in

a different random order for each trial. Shuffling the list reduces the utility of one possible organisational strategy that participants may use to structure their own recall, i.e. recalling the words in the order they are presented.

The hypothesis is that all patient groups will show impaired memory function and poorer subjective organisation. When controlling for the degree of organisation, it is expected that group differences in recall will no longer remain significant for euthymic patients, but will remain significant for depressed patients.

METHODS

PARTICIPANTS

For the verbal memory study, three different sets of participants (both patients and controls) were used for each of the different administration conditions of the verbal learning test. In total this included one sample of euthymic patients (n=62, with n=62 controls) who received standard administration, one sample of depressed patients (n=38, with n=23 controls) who received standard administration, and a further sample of euthymic patients (n=38, with n=27 controls) who received the shuffled administration.

1. DEPRESSED PATIENTS – STANDARD ADMINISTRATION

Patients aged 18 to 65 years with a diagnosis of bipolar disorder, confirmed using the Structured Clinical Interview for DSM-IV (SCID; First et al., 1995), were recruited from secondary and tertiary care services in North East of England. All were currently in a depressive episode (HAMD score >15). Patients were excluded if they met criteria for any other current axis I disorder, including anxiety disorder, schizophrenia or substance dependence/abuse. Illness characteristics, clinical ratings and medication history were determined by trained psychiatrists using full history, case-note and medication review

and standardised rating scales. All patients were receiving medication at the time of testing which had remained stable for a minimum of 4 weeks.

Healthy control subjects were recruited by advertisement and from hospital/university staff. All were physically healthy and had no personal or family history of psychiatric illness.

After a complete description of the study, written informed consent was obtained from all participants. The study was approved by Newcastle and North Tyneside Research Ethics Committee.

2. EUTHYMIC PATIENTS – STANDARD ADMINISTRATION

Sixty two euthymic patients with bipolar disorder were recruited from outpatient departments in the North East of England. Diagnosis was confirmed with the SCID-I/P (First et al., 1995) administered by a psychiatrist. All patients were euthymic for at least a month before testing (Hamilton Depression Rating scale(Hamilton, 1960) (HAM-D) ≤ 7 and Young Mania Rating Scale (Young et al., 1978) (YMRS) ≤ 7) which was verified prospectively. Exclusion criteria comprised: comorbid axis I diagnosis; significant neurological or medical condition; ECT within the last year (8 patients had received ECT previously); history of substance or alcohol misuse within the past 6 months, or current alcohol intake above 28 or 21 units/week for males or females respectively (4 patients met DSM-IV criteria for a previous history of alcohol or substance abuse); currently taking corticosteriods or anti-hypertensives.

Sixty two healthy volunteers were recruited from the region by advertisement. Controls were matched to patients on the basis of age, sex, handedness and premorbid IQ. Controls were interviewed with the SCID-I/NP and were excluded if they demonstrated a current or past psychiatric illness or had a family history of affective

disorder in a first-degree relative, or were currently receiving any medication (other than the oral contraceptive pill).

Participants gave written informed consent to participant and the study was approved by the Local Research Ethics Committee.

3. EUTHYMIC PATIENTS – SHUFFLED ADMINISTRATION AND WISCONSIN CARD SORTING TEST

The details of the clinical sample are the same as those described previously in Chapter 5, page 97.

MATERIALS

1. REY AUDITORY VERBAL LEARNING TEST (RAVLT) (REY, 1964)

STANDARD ADMINISTRATION

In this test the participant is read a 15-item list of unrelated words and is asked to recall as many items from the list as they can (trial 1). The same procedure is followed four further times, with the same list of words (list A) read to the participant in the same order and recall measured after each presentation of the list (trials 2-5). After the fifth presentation, a different 15-item word list, list B, is read once to the participant and recall is assessed. After this distractor list, the participant is asked to recall as many items as possible from list A without hearing the words again (trial 6). After a 30minute delay (which in each study was filled with non-verbal tasks), recall of list A is assessed again (trial 7), followed by a recognition trial. In the recognition trial, participants are shown 50 words on a page – all 15 from list A, all 15 from list B, and 20 words that were not on either list – and are asked for each word to indicate whether it was on the first list, the second list, or neither list.

There are a number of performance indices derived from this task (see Chapter 5, page 105, for further detail), but the indices of interest for the present chapter are recall on each of trials 1 to 7 of list A.

SHUFFLED ADMINISTRATION

In the standard administration of the task, list A is presented in the same order on trails 1-5. For the shuffled presentation, the word lists were presented in a random order on each presentation. To randomise the lists, each word was numbered from one to 15, and four different randomised orders were created for the numbers one to 15 using an online randomiser (http://www.randomizer.org/form.htm). The first list presentation was left in the same order as the standard administration to enable a comparison between all three samples on recall of this list. The random word lists were checked to assure there were no obvious strings of words in the same order, however they were not constrained to be entirely random. The justification behind this is that a set of lists containing no pairwise repetitions would result in organisation that is below-chance level if the participant recalls the items in perfect serial order (Murphy and Puff, 1982). Calculating the degree of organisation for the list-pairs as presented resulted in ARC' measures of 0.09, -0.07, 0.09, and -0.07 for the pairs A1-A2, A2-A3, A3-A4, and A4-A5 respectively. The same values for the SO measure were 0.19, 0.06, 0.19, 0.06. Each participant received the five randomised lists in the same order.

PROCEDURE

Although conducted as part of separate studies, in all cases the verbal learning test was administered as part of a larger battery of tests. The delay interval between trials 1-6 and trial 7 was filled with other cognitive tasks.

DATA ANALYSIS

To assess the level of subjective organisation, two measures were calculated – Pellegrino & Pellegrino's ARC' and Tulving's SO. The two measures were calculated for each adjacent pair of lists for each participant using an SPSS script using formulae (1) and (2) above (Kazen and Otani, 1996). This yielded six subjective organisation values for each participant, for each of two different measures (12 in total).

To explore whether participants recalled the words in a similar order to the presented list, a second set of subjective organisation measures were calculated contrasting the presentation order of the list with the subsequent recall of that list. For the participants receiving standard administration this second set of measures were calculated for all 7 recalls to investigate whether trials where no list is presented prior to recall (trials 6 and 7) are recalled in a similar order to the presented words. For those receiving the shuffled administration, this second set of measures were only calculated for the five trials on which a list was presented. It was not possible to calculate the presentation-recall organisation scores for trials 6 and 7 of the shuffled presentation as there was no obvious presentation order to use for the comparison.

Data were analysed using SPSS version 17.0. The mean age of the patient and control groups were compared using independent samples t-tests. The significance level was set to $p \leq 0.05$, with $p < 0.1$ representing a trend toward significance.

For the verbal memory task, analyses proceeded for each group individually as follows: firstly patient and control groups were compared using independent samples t-tests on recall of trials 1 to 7 in order to establish whether performance on the task differed. Subsequently, the patient and control groups were compared on the two subjective organisation measures using independent samples t-tests for each of the trial pairs from 1 to 7 (i.e. 6 trial pairs in total). In the next stage of the analysis, the patient

and control groups were compared on the subjective organisation measures calculated between the presentation list and the recall of that list. Finally, Pearson correlations were computed between the subjective organisation indices and recall for the patient and control groups separately, and ANCOVA were used to compare patient and control groups on recall of list p covarying for subjective organisation between list p and list p-1.

For t-tests, Levene's F-test was first used to identify instances of unequal variance, and if $p<0.05$ corrected p-values were reported. The data were screened for outliers in the same way described in Chapter 5, on page 115, and variables with notable outliers were analysed with non-parametric tests.

In an exploratory analysis to explore whether the subjective organisation indices reflect executive processes (or other cognitive processes), the data from the euthymic sample with shuffled administration were investigated further (this sample was chosen as subjective organisation indices are specifically designed to be calculated on recall output from shuffled lists). The ARC' and SO measures for each trial-pair were correlated with selected executive, psychomotor and non-list-learning verbal memory measures using pearson correlations. At least one measure from each executive function (according to the taxonomy used in chapter 5: fluency, inhibition, planning, mental manipulation, and set-shifting) was selected, as well as psychomotor measures and a verbal memory measure derived from a paragraph recall task. These additional measures were included to establish whether the subjective organisation measures relate broadly to several measures or specifically to executive measures. As an additional exploratory measure, total recall on the RAVLT was correlated with the same measures. In the case where subjective organisation measures do not correlate with executive measures, this additional measure was included to establish whether this is because recall does not correlate with executive function. The correlations were performed separately in

patients and controls to identify whether the two groups show a different pattern of interrelationships between the subjective organisation measures and the other cognitive measures.

Differences between the three control groups or the three patient groups were explored using one way ANOVA with group (control/patient group depressed sample, control/patient group euthymic with standard administration, control/patient group with shuffled administration) as the independent variable and recall or subjective organisation as the dependent variable. Post hoc LSD tests were used to follow-up differences (using a significance level of $p<0.05$).

To explore the convergent validity of the subjective organisation measures, within each sample the SO and ARC' measures were correlated with one another using Pearson correlations.

Effect sizes were calculated using Cohen's d where d=(control mean-patient mean)/pooled standard deviation. Effect sizes for indices where a higher score indicated worse performance (e.g. error or reaction time indices) were converted (multiplied by -1) such that a positive effect size always indicates a poorer performance by the patient group. Conventions for Cohen's d are that $0.2<d<0.5$ is a small effect, $0.5\leq d<0.8$ is a medium effect, and $d\geq 0.8$ is a large effect.

RESULTS

DEPRESSED PATIENTS – STANDARD ADMINISTRATION:

Performance on the memory task is reported in Table 6.13 below. Graphical representations of the recall and subjective organisation data are presented in Appendix 5, on page 317. Depressed patients with bipolar disorder showed significantly lower recall than controls for trials 1 to 5 (all $p<0.05$). Effect sizes indicated medium to large effects ($0.4 \leq d \leq 1.02$). Group differences were statistically non-significant or significant at trend levels only for trials 6 ($p=0.131$) and 7 ($p=0.09$).

Table 6.13: Average words recalled for trials 1-7 of the verbal learning task. Depressed patients, standard administration

	Control (n=23)		Bipolar (n=38)				
	mean	s.d.	mean	s.d.	d	t_{59}	p
Age	44.6	15.7	46.9	10.3	-0.18	0.69	0.493
Trial 1	6.1	1.59	4.7	1.3	1.02	-3.85	<0.001
Trial 2	8.1	2.27	7.0	1.8	0.54	-2.03	0.047
Trial 3	10.0	2.12	8.4	2.3	0.67	-2.52	0.014
Trial 4	10.8	2.41	9.2	2.4	0.69	-2.62	0.011
Trial 5	11.3	2.46	9.8	2.3	0.65	-2.47	0.017
Trial 6	9.2	3.69	7.9	3.0	0.40	-1.53	0.131
Trial 7	8.3	3.58	6.8	3.4	0.45	-1.72	0.090

Comparing subjective organisation between patients and controls indicates no significant differences between the groups on either measure for any trial (Table 6.14 below). On several trials, the degree of organisation in the patient group was higher than that in the control group, although not statistically significantly so. There is a relative increase in organisation between trials 6 and 7 (the two trials on which the list is not

presented to the participants before recall) in the control group for the ARC' measure (see Appendix 5, Figure A5.11, page 317). The same pattern is not evident in the patient group (and is not evident for either group on the SO measure).

Table 6.14: Subjective Organisation measures (ARC' and SO) for consecutive trial pairs. Depressed patients, standard administration

	Control (n=23)		Bipolar (n=38)				
	mean	s.d.	mean	s.d.	d	t_{59}	p
ARC' A1 vs A2	0.17	0.2	0.12	0.3	0.17	-0.73	0.466
ARC' A2 vs A3	0.10	0.2	0.17	0.2	-0.40	1.51	0.136
ARC' A3 vs A4	0.12	0.2	0.17	0.2	-0.25	846.5[a]	0.056
ARC' A4 vs A5	0.14	0.2	0.14	0.2	-0.01	0.05	0.964
ARC' A5 vs A6	0.12	0.1	0.14	0.2	-0.10	0.40	0.689
ARC' A6 vs A7	0.22	0.4	0.14	0.2	0.23	-0.89	0.378
SO A1 vs A2	0.30	0.2	0.29	0.2	0.08	-0.30	0.769
SO A2 vs A3	0.27	0.2	0.32	0.2	-0.28	1.06	0.292
SO A3 vs A4	0.24	0.2	0.32	0.2	-0.38	1.45	0.152
SO A4 vs A5	0.26	0.2	0.29	0.2	-0.15	0.55	0.585
SO A5 vs A6	0.27	0.1	0.31	0.2	-0.25	0.93	0.357
SO A6 vs A7	0.29	0.2	0.34	0.2	-0.22	0.83	0.411

[a] Mann-Whitney U

Investigating whether the groups organised recall output in line with the presented list showed that there was very little difference between the groups (see Table 6.15 below). For trial 1, the ARC' measure indicated that controls organised significantly more in line with the list than patients (p=0.042). In contrast, for trial 3 the SO measure indicated the opposite, that patients organised more in line with the list than controls, although this result was on the boundary of statistical significance (p=0.051).

Neither group in this sample showed evidence that recall output was increasingly organised in line with the list across trials. Although trials 3 to 5 showed higher organisation scores than trials 1 and 2 for both groups, the difference was small.

Table 6.15: Subjective organisation measures comparing organisation between the presented list and the participant's recall. Depressed patients, standard administration

	Control (n=23)		Bipolar (n=38)				
	mean	s.d.	mean	s.d.	d	t_{59}	p
ARC'							
Presentation 1 : Recall 1	0.09	0.1	0.05	0.1	0.55	-2.08	0.042
Presentation 1 : Recall 2	0.09	0.1	0.10	0.1	-0.12	0.44	0.660
Presentation 1 : Recall 3	0.12	0.1	0.15	0.1	-0.24	0.92	0.364
Presentation 1 : Recall 4	0.18	0.1	0.14	0.1	0.34	-1.29	0.202
Presentation 1 : Recall 5	0.16	0.1	0.15	0.2	0.07	-0.27	0.785
Presentation 1 : Recall 6	0.15	0.2	0.10	0.2	0.24	-0.92	0.360
Presentation 1 : Recall 7	0.10	0.2	0.08	0.1	0.13	-0.49	0.628
SO							
Presentation 1 : Recall 1	0.37	0.2	0.30	0.2	0.39	-1.47	0.146
Presentation 1 : Recall 2	0.26	0.2	0.32	0.1	-0.40	1.53	0.132
Presentation 1 : Recall 3	0.27	0.1	0.35	0.1	-0.53	1.99	0.051
Presentation 1 : Recall 4	0.30	0.2	0.29	0.1	0.07	-0.25	0.803
Presentation 1 : Recall 5	0.30	0.1	0.30	0.2	0.02	-0.06	0.952
Presentation 1 : Recall 6	0.31	0.2	0.33	0.2	-0.09	0.35	0.726
Presentation 1 : Recall 7	0.25	0.2	0.27	0.2	-0.13	0.51	0.612

The correlations between subjective organisation and number of words recalled are reported in Table 6.16 below. Each correlation represents organisation between lists p and p+1 correlated with recall on list p+1. For the ARC' measure, recall and subjective organisation were significantly correlated for almost all trials. In general, the correlations were stronger in the control sample than in the bipolar patients (although it should be noted that only one pair of correlations showed a statistically significant difference between the two groups. The correlation for ARC' between lists A6 and A7 and recall on list A7 was significantly higher in the control group than the patient group ($p<0.001$). None of the other correlation-pairs showed a statistically significant difference between

the two groups, all p>0.11). There was one anomalous result in the control sample – a negative correlation between organisation between trials 6 and 7 and recall on trial seven, indicating lower organisation was associated with better recall, although this did not reach statistical significance.

The relationship between the SO measure and recall indicated that, for both groups, this measure is less strongly related to number of words recalled. The correlations were lower for both groups than those reported with the ARC' measure, and only three were strong enough to reach statistical significance (see Table 6.16). However, when compared statistically, there were no statistically significant differences within each participant group for the correlations between ARC' and recall and SO and recall (with the exception of the correlation with recall on list A7, where the correlation in the control group was significantly higher between recall on A7 and the SO measure than it was between recall on A7 and the ARC' measure (p=0.032)).

Table 6.16: Pearson correlations between subjective organisation and recall. Depressed patients, standard administration

	Bipolar (n=38)		Control (n=23)	
	r	p	r	p
ARC'				
A1:A2 - A2	0.38	0.018	0.13	0.544
A2:A3 - A3	0.21	0.197	0.47	0.025
A3:A4 - A4	0.32	0.050	0.61	0.002
A4:A5 - A5	0.27	0.096	0.51	0.013
A5:A6 - A6	0.57	<0.001	0.60	0.002
A6:A7 - A7	0.55	<0.001	-0.36	0.090
SO				
A1:A2 - A2	0.09	0.606	-0.09	0.696
A2:A3 - A3	-0.06	0.745	0.31	0.150
A3:A4 - A4	0.11	0.520	0.50	0.016
A4:A5 - A5	0.12	0.458	0.24	0.264
A5:A6 - A6	0.27	0.103	0.47	0.024
A6:A7 - A7	0.39	0.016	0.29	0.178

The final stage of analysis comparing group differences covarying for subjective organisation is reported in Table 6.17 below. Each ANCOVA included one between groups factor as a fixed factor (patient/control) and one covariate (subjective organisation – either ARC' or SO – between lists p and p+1), with recall on list p+1 as the dependent variable. For the ARC' measure, the covariate was significant for all trials ($p<0.05$) except trial 7 ($p=0.686$). Significant differences between groups on recall remained for trials 3 to 5 ($p<0.05$), but the significance of the difference on trial 2 was reduced to trend levels ($p=0.062$; it was previously $p=0.047$ in the comparison without a covariate). The inclusion of the covariate allowed the emergence of a significant difference on trial 6 ($p=0.042$) that was previously non-significant in straightforward analyses. Overall, covarying for ARC' did not reduce or remove group differences on verbal recall.

A very similar picture can be seen for the SO measure in that covarying for SO did not remove group differences on recall. SO was only a significant (or near-significant) covariate for trials 4, 6 and 7.

Table 6.17: Results of two separate ANCOVAs comparing performance between groups on the list-learning task covarying for 1) ARC' or 2) SO. Depressed patients, standard administration

	ANCOVA 1				ANCOVA 2				d	d	d
	ARC'		Group		SO		Group		without covariate	covary ARC'	covary SO
	$F_{1,58}$	p	$F_{1,58}$	p	$F_{1,58}$	p	$F_{1,58}$	p			
Trial 1	-	-	-	-	-	-	-	-	-	-	-
Trial 2	5.20	0.026	3.62	0.062	0.02	0.882	4.02	0.05	0.54	0.50	0.53
Trial 3	6.11	0.016	9.35	0.003	0.27	0.606	6.53	0.013	0.67	0.80	0.68
Trial 4	13.10	0.001	10.86	0.002	3.30	0.075	8.74	0.004	0.69	0.87	0.79
Trial 5	9.54	0.003	7.06	0.01	1.75	0.192	6.60	0.013	0.65	0.70	0.68
Trial 6	26.56	<0.001	4.32	0.042	7.13	0.01	3.67	0.06	0.40	0.55	0.51
Trial 7	0.17	0.686	2.73	0.104	8.18	0.006	4.49	0.038	0.45	0.44	0.56

Considering these findings for the depressed patients together, there is evidence of significant impairment in verbal recall with moderate to large effect sizes ($0.45 \leq d \leq 1.02$). There was no evidence of impairment in subjective organisation. The ARC' measure correlated significantly and positively with recall in both groups, however between-group differences in recall were broadly unchanged by covarying for the degree of subjective organisation.

EUTHYMIC PATIENTS – STANDARD ADMINISTRATION:

Performance on the verbal memory task is reported in Table 6.18 below. Graphical representations of the recall and subjective organisation data are presented in Appendix 5, on page 317. Patients were significantly impaired relative to controls for each recall trial (all $p<0.05$) apart from trial 1 ($p=0.25$). Effect sizes were small to medium, becoming larger as the task progressed. In general the effect sizes were smaller than those reported in the depressed sample described above.

Table 6.18: Average words recalled for trials 1-7 of the verbal learning task. Euthymic patients, standard administration

	Control (n=62)		Bipolar (n=62)				
	mean	s.d.	mean	s.d.	d	t_{122}	p
Age	45.3	9.1	44.2	8.6	0.12	0.68	0.498
Trial 1	6.5	1.6	6.1	1.5	0.20	1.17	0.245
Trial 2	9.2	2.1	8.3	1.8	0.47	2.55	0.012
Trial 3	11.1	2.2	10.2	2.2	0.40	2.32	0.022
Trial 4	12.0	2.0	11.0	2.4	0.45	2.69	0.008
Trial 5	12.7	1.8	11.6	2.5	0.47	2.65	0.009
Trial 6	10.7	2.3	9.4	2.9	0.50	2.90	0.004
Trial 7	10.5	2.7	9.1	2.9	0.52	2.90	0.004

The results of subjective organisation scores are shown in Table 6.19 below. In both groups the level of organisation gradually increases across trials 1-5 when the list is presented before recall. This can be seen in both measures of subjective organisation

and is more pronounced in the control group than the patient group. Organisation drops between trials 5 and 6 (the lists which are separated by a distractor list) and, in the control group especially, increases markedly between trials 6 and 7 (see Appendix 5, Figure A5.13, page 318). Organisation in the control group is at its highest between trials 6 and 7, and in the patient group is almost at its highest. The same pattern was noted in the control group described above.

The only between-group difference that is statistically significant at conventional levels is the ARC' measure calculated between trials 6 and 7 ($p=0.023$) – the two recall trials where the list of words is not presented beforehand. Patients exhibited significantly lower subjective organisation than controls between these two trials. Differences in SO for the same trial-pair showed a trend towards statistical significance ($p=0.096$), as did organisation between lists 3 and 4 for the ARC' measure ($p=0.068$).

In general, the control group in this sample showed a higher level of organisation than the control group recruited for the previous depressed sample, however the two patient groups organised to a very similar extent. Testing this formally (using LSD post-hoc tests following one way ANOVA with three groups (the three control groups from each of the studies)) indicated that the control group for the depressed patients organised significantly less than the control group from the euthymic study with standard administration for 6 out of 12 comparisons (ARC': A2-A3, $p=0.033$; A3-A4, $p=0.027$; A4-A5, $p=0.041$; SO: A3-A4, $p=0.021$; A4-A5, $p=0.016$; A6-A7, $p=0.028$). The patient groups showed no significant differences on any of the 12 indices (all $p>0.05$).

Table 6.19: Subjective Organisation measures (ARC' and SO) for consecutive trial pairs. Euthymic patients, standard administration

	Control (n=62)		Bipolar (n=62)				
	mean	s.d.	mean	s.d.	d	t_{122}	p
ARC' A1 vs A2	0.10	0.2	0.10	0.3	0.01	0.03	0.976
ARC' A2 vs A3	0.20	0.2	0.17	0.2	0.16	0.88	0.380
ARC' A3 vs A4	0.22	0.2	0.16	0.2	0.33	1.85	0.068
ARC' A4 vs A5	0.23	0.2	0.21	0.2	0.13	0.74	0.464
ARC' A5 vs A6	0.18	0.2	0.15	0.2	0.16	0.87	0.385
ARC' A6 vs A7	0.26	0.2	0.17	0.2	0.41	2.31	0.023
SO A1 vs A2	0.24	0.2	0.29	0.2	-0.23	-1.26	0.211
SO A2 vs A3	0.31	0.2	0.31	0.2	0.05	0.27	0.789
SO A3 vs A4	0.34	0.2	0.30	0.2	0.22	1.21	0.230
SO A4 vs A5	0.36	0.2	0.34	0.2	0.09	0.52	0.607
SO A5 vs A6	0.32	0.2	0.31	0.2	0.08	0.46	0.650
SO A6 vs A7	0.39	0.2	0.34	0.2	0.30	1.68	0.096

The subjective organisation indices comparing whether recall output mirrored the order of the presented list are shown in Table 6.20 below. The degree of organisation gradually increases across trials 1-5 for both groups, and this difference is more marked for the patient group than the control group. Numerically, the data for the ARC' measure indicate that across all trials (except trial 5), the similarity between the order of presentation and the order of recall is higher in the control group than it is in the patient group. For the ARC' measure this difference is statistically significant for trial 1 and trial 7 (p=0.004, p=0.017 respectively) and showed a trend towards statistical significance for trials 4 and 6 (p=0.053, p=0.062 respectively).

The pattern for the SO measure is slightly different. Controls have a relatively flat curve showing little change across all 7 trials. In contrast, patients show a gradual increase in the level of organisation from below that of controls on trial 1 to above that of controls by trial 5. For trials 6 and 7 patients experience a larger drop in organisation

than controls, and indeed the only significant difference between patients and controls on this measure is on trial 7 (p=0.047). Patients did not organise their output in line with the presented list to the same degree as controls.

Table 6.21 below shows the correlations between subjective organisation and recall. The ARC' measure correlates significantly with recall for each trial pair in both patients and controls (all p<0.05) apart from the correlation with recall on A2 in the patient group, which is significant at trend levels only (p=0.054). However, in all instances the correlations are higher in control participants than in the patients and indicate that greater subjective organisation is associated with better recall. It must be noted, however, that correlations in the two groups are not statistically significantly different when contrasted directly (all p>0.3).

Table 6.20: Subjective organisation measures comparing organisation between the presented list and the participant's recall. Euthymic patients, standard administration

	Control (n=62)		Bipolar (n=62)				
	mean	s.d.	mean	S.d.	d	t_{122}	p
ARC'							
Presentation 1 : Recall 1	0.08	0.1	0.04	0.1	0.53	1268.0[a]	0.001
Presentation 1 : Recall 2	0.14	0.1	0.12	0.1	0.18	0.99	0.326
Presentation 1 : Recall 3	0.19	0.1	0.17	0.2	0.12	0.66	0.510
Presentation 1 : Recall 4	0.23	0.2	0.17	0.1	0.35	1.96	0.053
Presentation 1 : Recall 5	0.23	0.2	0.24	0.2	-0.04	-0.21	0.834
Presentation 1 : Recall 6	0.17	0.2	0.11	0.1	0.34	1.88	0.062
Presentation 1 : Recall 7	0.16	0.2	0.09	0.1	0.43	2.42	0.017
SO							
Presentation 1 : Recall 1	0.30	0.2	0.26	0.1	0.24	1.33	0.187
Presentation 1 : Recall 2	0.30	0.1	0.31	0.1	-0.09	-0.52	0.607
Presentation 1 : Recall 3	0.33	0.1	0.32	0.2	0.07	0.42	0.679
Presentation 1 : Recall 4	0.34	0.2	0.31	0.2	0.18	1.00	0.318
Presentation 1 : Recall 5	0.34	0.2	0.37	0.2	-0.17	-0.92	0.259
Presentation 1 : Recall 6	0.32	0.2	0.28	0.2	0.24	1.35	0.180
Presentation 1 : Recall 7	0.32	0.2	0.26	0.2	0.36	2.00	0.047

[a] Mann-Whitney U

Table 6.21: Pearson correlations between subjective organisation and recall. Euthymic patients, standard administration

	Bipolar (n=62)		Control (n=62)	
	r	p	r	p
ARC'				
A1:A2 - A2	0.24	0.054	0.41	0.001
A2:A3 - A3	0.40	0.002	0.44	<0.001
A3:A4 - A4	0.37	0.004	0.49	<0.001
A4:A5 - A5	0.27	0.037	0.30	0.017
A5:A6 - A6	0.55	<0.001	0.59	<0.001
A6:A7 - A7	0.30	0.016	0.45	<0.001
SO				
A1:A2 - A2	0.14	0.287	0.34	0.007
A2:A3 - A3	0.25	0.048	0.40	0.001
A3:A4 - A4	0.12	0.352	0.46	<0.001
A4:A5 - A5	0.13	0.321	0.20	0.116
A5:A6 - A6	0.38	0.002	0.49	<0.001
A6:A7 - A7	0.19	0.136	0.38	0.003

For the SO measure, the broad pattern is the same, but overall the correlations are weaker (although not statistically significantly so for either participant group; there is no statistically significant difference between the strength of the correlations between ARC' and recall or SO and recall in either group, all p>0.15). The correlations are larger in the control group than in the patient group, but again these differences do not reach statistical significance in all but one instance (the correlation between SO on trials A3:A4 and recall on A4 is significantly higher in the control group (p=0.041); all other p>0.24). All correlations bar one are significant in control participants, however only two reach significance in the patient sample. The positive correlations indicate that, as was the case with ARC', a greater degree of organisation is associated with a larger number of words recalled. This effect is notably stronger in the control group than the patient group.

Considering the above series of results together – the fact that patients generally show lower levels of subjective organisation than controls and that subjective organisation correlates positively with recall in both groups – it could be expected that reinvestigating group differences in recall whilst covarying for the level of subjective organisation would moderate or remove differences. The results of the ANCOVAs are reported below in Table 6.22. The results indicate that although the subjective organisation measure was a statistically significant covariate in all cases (apart from trial 5 for the SO measure which showed a trend towards significance (p=0.083)), group differences still remained on all trials irrespective of which subjective organisation measure was used as the covariate (although note that differences on trial 4 only reached a trend level of significance when ARC' was included as a covariate (p=0.051)).

Table 6.22: Results of two separate ANCOVAs comparing performance between groups on the list-learning task covarying for 1) ARC' or 2) SO. Euthymic patients, standard administration

	ANCOVA 1				ANCOVA 2				d	d	d
	ARC'		group		SO		Group		without covariate	covary ARC'	covary SO
	$F_{1,119}$	p	$F_{1,119}$	p	$F_{1,119}$	p	$F_{1,119}$	p			
Trial 1	-	-	-	-	-	-	-	-	-	-	-
Trial 2	13.64	<0.001	7.018	0.009	6.95	0.009	8.497	0.004	0.47	0.48	0.53
Trial 3	25.58	<0.001	3.96	0.049	14.683	<0.001	4.912	0.029	0.40	0.36	0.40
Trial 4	26.078	<0.001	3.88	0.051	10.61	0.001	5.415	0.022	0.45	0.36	0.42
Trial 5	10.118	0.002	6.407	0.013	3.05	0.083	6.703	0.011	0.47	0.45	0.47
Trial 6	55.24	<0.001	8.332	0.005	26.507	<0.001	8.83	0.004	0.50	0.52	0.53
Trial 7	19.261	<0.001	4.596	0.034	10.312	0.002	6.208	0.014	0.52	0.39	0.45

Overall, these results have demonstrated that there is evidence that subjective organisation is lower in euthymic patients with bipolar disorder relative to controls, and that subjective organisation relates significantly to verbal recall, however the differences in subjective organisation are not sufficient to account for the entire verbal learning deficit evident in patients.

EUTHYMIC PATIENTS – SHUFFLED ADMINISTRATION:

Group differences in recall are shown in Table 6.23 below. Graphical representations of the recall and subjective organisation data are presented in Appendix 5, on page 319. Differences at conventional levels were evident on trials 1 and 2 (p=0.028 and p=0.048 respectively), and at trend levels on trials 3 and 6 (p=0.073 and p=0.099 respectively), all indicating worse performance by the patient group. The pattern of findings differs for this sample in comparison to the previous two. There are fewer significant differences and the effect sizes decrease across trials rather than increase or stay the same (with trial 6 as the exception). Comparing the performance of the present two groups with the performance of the other two samples, it is clear that shuffling the order of the lists has negatively impacted performance for both groups. However, this effect appears to be more pronounced for the control group, especially on later trials. Calculating the recall of the present control sample as a percentage of the recall of the euthymic sample who received standard administration indicates that on average – ignoring other possible differences between groups – recall with the shuffled list is between 81% and 99% of recall with standard administration. However, contrasting the two euthymic patient samples in the same manner shows that recall in the shuffled administration condition did not reach quite as low a proportion of recall under standard conditions (range 88% to 94%). Comparing the two control groups with one another on recall performance (using post-hoc LSD tests following one way ANOVA with three groups (the three control groups) as the independent variable and recall on each trial of the RAVLT as the dependent variable) indicated significant differences between the two on trials A4 (p=0.004), A5 (p=0.001), A6 (p=0.032), and A7 (p=0.003). In contrast, the two patient groups showed statistically significant differences on trial A1 only (p=0.036).

Table 6.23: Average words recalled for trials 1-7 of the verbal learning task. Euthymic patients, shuffled administration

	Control (n=28)		Bipolar (n=38)				
	mean	s.d.	mean	s.d.	d	t_{64}	p
Age	46.5	10.8	44.9	12.7	0.13	0.54	0.592
Trial 1	6.4	1.6	5.5	1.6	0.57	2.25	0.028
Trial 2	8.7	1.7	7.8	1.9	0.52	2.01	0.048
Trial 3	10.2	1.7	9.4	1.9	0.47	1.82	0.073
Trial 4	10.7	2.1	9.9	2.3	0.37	1.43	0.158
Trial 5	11.2	1.6	10.6	2.3	0.28	1.13	0.265
Trial 6	9.4	2.4	8.3	2.8	0.42	1.67	0.099
Trial 7	8.6	2.5	8.1	3.2	0.17	0.68	0.499

Looking at the subjective organisation measures in Table 6.24 below, similar to the findings for the depressed patients there are few differences between the groups and several trials on which subjective organisation was higher in the patient than the control group. The only difference to reach statistical significance was the degree of organisation between trials 6 and 7 for both organisation indices (ARC' p=0.014, SO p=0.041), indicating a lower degree of organisation in the patient group for this trial pair.

Table 6.24: Subjective Organisation measures (ARC' and SO) for consecutive trial pairs. Euthymic patients, shuffled administration

	Control (n=28)		Bipolar (n=38)				
	mean	s.d.	mean	s.d.	d	t_{64}	p
ARC' A1 vs A2	0.05	0.1	0.09	0.2	-0.25	-1.01	0.318
ARC' A2 vs A3	0.04	0.1	0.06	0.2	-0.09	-0.36	0.717
ARC' A3 vs A4	0.00	0.1	0.02	0.1	-0.27	-1.09	0.281
ARC' A4 vs A5	0.05	0.2	0.07	0.2	-0.10	-0.38	0.705
ARC' A5 vs A6	0.08	0.1	0.06	0.1	0.14	0.57	0.569
ARC' A6 vs A7	0.24	0.2	0.11	0.2	0.68	2.57	0.014
SO A1 vs A2	0.19	0.2	0.20	0.2	-0.04	-0.16	0.874
SO A2 vs A3	0.17	0.1	0.17	0.1	0.03	0.10	0.920
SO A3 vs A4	0.13	0.1	0.15	0.1	-0.23	-0.95	0.343
SO A4 vs A5	0.17	0.1	0.18	0.1	-0.13	-0.53	0.596
SO A5 vs A6	0.21	0.1	0.20	0.1	0.01	0.03	0.974
SO A6 vs A7	0.40	0.2	0.31	0.2	0.52	2.09	0.041

There were no significant differences between patients and controls in the extent to which their recall output mirrored the list (see Table 6.25 below). In general, levels of organisation relative to the list were low and showed little change across trials.

The correlations between subjective organisation and recall are reported in Table 6.26 below. In contrast to the two samples reported above, the relationship between subjective organisation and recall is much weaker in the present sample – especially for earlier recall trials. However, in line with earlier findings, the correlations are stronger in the control group and a greater number are sufficiently strong to reach statistical significance than is the case in the patient group (see Table 6.26). Several of the relationships are significantly stronger in the control participants than in the patient group (ARC' A1:A2-Recall A2, p=0.009; ARC' A5:A6-Recall A6, p=0.040; AO A1:A2-Recall

A2, p=0.051; SO A5:A6-Recall A6, p=0.009; SO A6:A7-Recall A7, p=0.002; all other p>0.23). Most notably, the correlations are strongest for the final recall trials, the point at which participants are most familiar with the list-items and have had maximum opportunity to structure their recall output consistently. In the patient group, several of the correlations are negative, indicating lesser organisation was associated with better recall.

Table 6.25: Subjective organisation measures comparing organisation between the presented list and the participant's recall. Euthymic patients, shuffled administration

	Control (n=28)		Bipolar (n=38)				
	mean	**s.d.**	**mean**	**s.d.**	**d**	**t_{64}**	**p**
ARC'							
Presentation 1 : Recall 1	0.07	0.1	0.06	0.1	0.11	0.45	0.654
Presentation 1 : Recall 2	0.12	0.1	0.10	0.1	0.23	0.92	0.363
Presentation 1 : Recall 3	0.08	0.1	0.09	0.1	-0.09	-0.37	0.711
Presentation 1 : Recall 4	0.10	0.1	0.08	0.1	0.21	0.82	0.419
Presentation 1 : Recall 5	0.15	0.1	0.11	0.1	0.34	1.35	0.183
SO							
Presentation 1 : Recall 1	0.31	0.2	0.35	0.2	-0.25	-0.99	0.325
Presentation 1 : Recall 2	0.30	0.2	0.28	0.1	0.13	0.51	0.612
Presentation 1 : Recall 3	0.21	0.1	0.23	0.1	-0.15	-0.60	0.553
Presentation 1 : Recall 4	0.23	0.1	0.23	0.1	0.05	0.19	0.853
Presentation 1 : Recall 5	0.27	0.1	0.23	0.1	0.28	1.11	0.270

Table 6.26: Pearson correlations between subjective organisation and recall. Euthymic patients, shuffled administration

	Bipolar (n=38)		Control (n=28)	
	r	p	r	p
ARC'				
A1:A2 - A2	-0.29	0.10	0.37	0.06
A2:A3 - A3	-0.09	0.62	0.22	0.27
A3:A4 - A4	-0.08	0.65	0.22	0.28
A4:A5 - A5	0.22	0.20	0.34	0.08
A5:A6 - A6	0.17	0.31	0.61	0.001
A6:A7 - A7	0.35	0.03	0.58	0.001
SO				
A1:A2 - A2	-0.21	0.24	0.29	0.15
A2:A3 - A3	-0.11	0.54	0.18	0.38
A3:A4 - A4	-0.01	0.97	0.1	0.61
A4:A5 - A5	0.25	0.51	0.43	0.03
A5:A6 - A6	0.13	0.44	0.67	<0.001
A6:A7 - A7	-0.37	0.02	0.41	0.03

Re-examining the between-group differences in recall covarying for subjective organisation shows that the group differences are affected marginally. When covarying for ARC', the significant difference on trial 2 is reduced to trend-level significance only, but the marginal difference on trial 3 remains. This is despite the fact that ARC' is not a significant covariate for these trials. For trials 5 to 7, the covariate is significant, however there remain no significant differences between the two groups.

The pattern of between-group differences in recall remains the same when correcting for SO. The covariate is significant for trials 5 and 6, but has little impact on the between group differences in recall.

Table 6.27: Results of two separate ANCOVAs comparing performance between groups on the list-learning task covarying for 1) ARC' or 2) SO. Euthymic patients, shuffled administration

	ANCOVA 1				ANCOVA 2				d	d	d
	ARC'		group		SO		Group		without covariate	covary ARC'	covary SO
	F	p	F	p	F	p	F	p			
Trial 1	-	-	-	-	-	-	-	-	-	-	-
Trial 2	0.079	0.780	3.194	0.079	0.003	0.958	3.992	0.050	0.52	0.46	0.40
Trial 3	0.011	0.916	3.285	0.075	0.002	0.961	3.269	0.076	0.47	0.46	0.38
Trial 4	0.086	0.770	2.078	0.155	0.049	0.826	2.054	0.157	0.37	0.37	0.28
Trial 5	4.314	0.042	1.477	0.229	6.154	0.016	1.568	0.215	0.28	0.31	0.27
Trial 6	9.020	0.004	2.417	0.125	9.159	0.004	3.108	0.083	0.42	0.39	0.44
Trial 7	14.672	<0.001	0.266	0.608	0.193	0.662	0.584	0.447	0.17	-0.13	0.19

Overall, these results have demonstrated that the difference between patients and controls in verbal recall is less when a shuffled list-administration is used. There was no evidence that subjective organisation is lower in euthymic patients when the list is shuffled, apart from between the two trials on which no list is presented (trials 6 and 7). Subjective organisation showed a very weak relationship with recall and did not account for the entire verbal learning deficit evident in patients.

EXPLORATORY ANALYSIS OF THE RELATIONSHIP BETWEEN SUBJECTIVE ORGANISATION MEASURES AND EXECUTIVE FUNCTION

The correlations between the subjective organisation indices and other measures in the test battery are shown separately for the patients and controls in Table 6.28 (page 166) and Table 6.29 (page 167) respectively.

In the patient sample, there were several significant correlations between the subjective organisation indices and other measures from the battery. At least one subjective organisation measure correlated significantly with an executive measure from each of the identified executive domains (fluency, planning, mental manipulation and inhibition), apart from set-shifting. Correlations with psychomotor measures were also significant. However, the signs of the correlations were not consistent (e.g. category B

errors on the Hayling Sentence Completion Test correlated positively with ARC' between A4:A5, but negatively with ARC' between A5:A6). Also, the number of significant correlations (10 out of 90 for the ARC' measures and 6 out of 90 for the SO measures) was only marginally greater than the number that would be expected by chance (especially for the SO measures). Total recall showed a significant relationship with verbal fluency measures, a planning measure, a set-shifting measure and psychomotor speed.

In the control group there were fewer significant correlations between subjective organisation and other cognitive measures. Four correlations out of 180 reached statistical significance and indicated a positive relationship between subjective organisation and verbal fluency, inhibition and planning. Verbal recall showed a different pattern of relationships with the other measures in the control group than was seen in the patient group. Recall correlated significantly with verbal fluency, set-shifting and paragraph recall.

Table 6.28: Pearson correlations between subjective organisation measures, total recall on the list-learning test and other cognitive measures. Correlations for the patient group only (n=37)

	Bipolar Patients	ARC'						SO						RAVLT
		1:2	2:3	3:4	4:5	5:6	6:7	1:2	2:3	3:4	4:5	5:6	6:7	Tot 1-5
Flu-ency	Initial Letter Fluency	-0.39*	0.00	0.14	-0.06	0.08	0.18	-0.32‡	-0.04	0.16	-0.04	0.13	-0.08	0.43**
	Category Fluency	-0.06	0.11	0.22	0.12	0.25	0.49**	-0.06	0.06	0.05	0.18	0.30‡	0.06	0.51**
Inhib-ition	HSCT A errors	-0.01	0.18	-0.19	0.45**	-0.20	-0.01	-0.10	0.19	-0.21	0.36*	-0.24	0.14	-0.04
	HSCT B errors	-0.19	0.09	-0.13	0.35*	-0.46**	-0.22	-0.24	0.14	-0.05	0.31‡	-0.41*	-0.15	-0.05
	Stroop colour-word errors	-0.16	-0.16	-0.05	0.01	-0.15	-0.33‡	-0.19	-0.17	0.11	-0.02	-0.25	-0.16	-0.06
Plan ning	Zoo Map version 1 raw score	0.17	-0.04	-0.13	0.26	0.19	0.33‡	0.06	0.06	-0.15	0.22	0.28	0.14	0.21
	SOPT total errors	-0.07	0.44**	0.28	-0.22	-0.11	-0.02	-0.10	0.43**	0.30	-0.22	-0.06	0.07	-0.38*
M M	RDS span	-0.13	-0.23	-0.12	-0.01	0.40*	0.33*	-0.11	-0.25	-0.15	-0.06	0.43*	-0.05	0.20
Set-shifting	WCST perseverative errors	0.10	0.07	0.18	-0.03	-0.08	0.02	0.05	0.13	0.26	0.03	-0.06	0.21	-0.23
	WCST categories	0.08	-0.06	-0.09	-0.03	0.09	0.04	0.10	-0.11	-0.16	-0.10	0.08	-0.22	0.17
	TMTb time	0.01	0.08	-0.26	0.05	-0.11	-0.31‡	-0.12	0.10	-0.23	0.04	-0.19	0.11	-0.36*
Psycho-motor	DSST attempt 1	0.08	0.02	0.38*	0.06	0.20	0.53**	0.14	-0.06	0.30‡	0.11	0.34*	0.05	0.44**
	Learning Capacity slope	0.17	-0.06	-0.11	0.29	0.05	-0.12	0.27	-0.06	-0.14	0.25	0.10	-0.03	-0.18
	TMTa time	-0.05	0.32‡	-0.23	0.05	0.09	-0.14	-0.19	0.48**	-0.18	0.07	0.01	0.11	-0.35*
	Logical Passages Story units	0.08	-0.03	0.09	0.22	0.13	0.23	0.03	0.07	0.05	0.20	-0.01	-0.16	0.22

MM, Mental manipulation; HSCT, Hayling Sentence Completion Test; SOPT, Self-Ordered Pointing Test; RDS, Reverse Digit Span; WCST, Wisconsin Card Sorting Test; DSST, Digit Symbol Subsitution Test ; TMT, Trail Making Test

‡ 0.5<p<0.1, *p<0.05, ** p<0.01

Table 6.29: Pearson correlations between subjective organisation measures, total recall on the list-learning test and other cognitive measures. Correlations for the control group only (n=28)

Controls		ARC'						SO						RAVLT
		1:2	2:3	3:4	4:5	5:6	6:7	1:2	2:3	3:4	4:5	5:6	6:7	Tot 1-5
Flu-ency	Initial Letter Fluency	0.11	-0.15	-0.02	-0.07	-0.03	-0.36‡	0.22	-0.05	0.02	0.02	-0.16	-0.12	0.31
	Category Fluency	-0.14	-0.07	0.13	0.37‡	0.40*	0.17	-0.07	-0.03	0.10	0.40*	0.25	0.26	0.47*
Inhib-ition	HSCT A errors	-0.28	0.09	0.02	0.01	0.05	-0.05	-0.30	0.05	0.01	0.06	-0.05	-0.19	0.09
	HSCT B errors	-0.25	-0.06	-0.05	-0.20	-0.19	-0.19	-0.22	-0.06	0.12	-0.24	-0.16	-0.30	0.11
	Stroop colour-word errors	-0.33‡	-0.26	-0.01	-0.14	0.34‡	0.13	-0.39*	-0.27	0.12	-0.13	0.36	0.12	-0.19
Plan ning	Zoo Map version 1 raw score	-0.06	0.12	-0.33‡	-0.07	0.05	0.01	-0.07	0.14	-0.19	-0.06	0.00	0.00	0.08
	SOPT total errors	-0.17	-0.43*	0.06	0.11	-0.04	0.05	0.01	-0.21	0.00	0.03	0.14	0.10	-0.29
M M	RDS span	0.00	0.13	-0.11	-0.11	0.15	-0.34‡	-0.08	0.00	-0.14	-0.05	-0.02	-0.31	0.16
Set-shifting	WCST perseverative errors	0.12	0.03	-0.06	0.24	-0.23	-0.14	-0.02	0.04	-0.08	0.20	-0.15	-0.16	-0.26
	WCST categories	-0.11	-0.06	-0.02	-0.08	0.22	0.21	0.01	-0.01	-0.04	-0.07	0.18	0.16	0.39*
	TMTb time	0.26	0.31	0.07	-0.26	-0.04	0.08	0.33‡	0.17	0.04	-0.23	-0.03	-0.03	0.05
Psycho-motor	DSST attempt 1	0.23	-0.08	0.01	0.11	0.12	-0.16	0.11	-0.09	-0.09	0.06	0.13	-0.27	0.29
	Learning Capacity slope	-0.16	0.25	0.13	0.10	0.22	0.22	-0.29	0.28	0.16	0.12	0.18	0.23	0.09
	TMTa time	-0.22	0.04	0.09	-0.30	0.05	0.04	-0.23	-0.05	0.25	-0.24	-0.03	0.06	-0.07
	Logical Passages Story units	-0.07	0.13	0.07	0.07	0.20	0.29	-0.08	-0.01	0.10	0.08	0.07	0.27	0.42*

MM, Mental Manipulation; HSCT, Hayling Sentence Completion Test; SOPT, Self-Ordered Pointing Test; RDS, Reverse Digit Span; WCST, Wisconsin Card Sorting Test; DSST, Digit Symbol Subsitution Test ; TMT, Trail Making Test

‡ 0.5<p<0.1, *p<0.05, ** p<0.01

RELATIONSHIP BETWEEN SUBJECTIVE ORGANISATION INDICES

To explore the convergent validity of the two subjective organisation indices, the correlations between them are presented in Table 6.30. The two indices generally correlate very highly with one another, with almost all of the correlations exceeding r=0.7 in each of the three samples (with the exception of a single correlation – the correlation between ARC' and SO between trials A6 and A7 in the depressed sample which was only r=0.41). This indicates the two measures show a high degree of convergent validity.

Table 6.30: Pearson correlations between the ARC' and SO subjective organisation indices reported for each sample separately.

	Depressed (n=61)[a]	Euthymic Standard (n=124)[a]	Euthymic Shuffled (n=66)[a]
ARC' A1:A2 - SO A1:A2	0.73	0.79	0.77[c]
ARC' A2:A3 - SO A2:A3	0.71	0.90[b]	0.87
ARC' A3:A4 - SO A3:A4	0.84	0.94 [b]	0.84
ARC' A4:A5 - SO A4:A5	0.86	0.94	0.94
ARC' A5:A6 - SO A5:A6	0.85	0.93	0.86
ARC' A6:A7 - SO A6:A7	0.41	0.86	0.71

[a] All p<0.001
[b] n=123
[c] n=65

DISCUSSION

In this comparison of verbal recall and subjective organisation in depressed and euthymic patients with bipolar disorder, the results identified significant differences between recall on at least some trials of the list-learning test in all patient groups. Depressed patients who had received the standard administration showed the largest degree of impairment compared to their control group, with effect sizes between d=0.40

to d=1.02 and statistically significant differences on 5 trials out of 7. The euthymic patients who received the standard administration again showed extensive impairment, with statistically significant differences on 6 out of the 7 trials, but effect sizes covered a smaller range from d=0.20 to d=0.52. The euthymic patients who received the shuffled administration showed fewer statistically significant differences, with statistically significant differences on the first two trials only, but the effect sizes for the differences were of a similar magnitude to the other euthymic sample (d=0.17 to d=0.57). The pattern of impairment was very different across the three samples. Euthymic patients with standard administration showed a gradual increase in effect size across trials (i.e. a relative worsening), whereas the depressed patients showed a large initial impairment which was then relatively stable for the remaining immediate memory trials, and which was less pronounced by the delayed-recall trials (6 and 7). In contrast, euthymic patients with the shuffled list showed a lesser degree of impairment across trials, with the effect size becoming smaller on each additional trial (with the exception of trial 6).

With regard to subjective organisation, the two euthymic samples showed significantly lower subjective organisation than controls on only one trial-pair – organisation between trials 6 and 7 (both trials when the list is not presented). Depressed patients showed no differences in organisation relative to controls. Furthermore, neither the depressed patients nor the euthymic patients with the shuffled presentation showed any significant differences in how they ordered the list in relation to how the list was presented. The euthymic patients with standard administration showed significantly lower organisation relative to list-order than their control group. However, for both patients and controls, the subjective organisation indices calculated between consecutive recall outputs exceeded that calculated between list presentation and recall, indicating that factors other than the serial order of list items were being used to structure recall.

All patient and control groups showed a relationship between subjective organisation and recall for at least some of the trials, with the direction of the relationship generally indicating that higher levels of organisation were associated with better recall. However, analyses comparing recall between groups whilst covarying for the degree of subjective organisation were virtually unchanged relative to uncontrolled comparisons. These results only partially support the hypothesis – group differences in recall were predicted to remain in depressed patients after subjective organisation was accounted for, but to be explained by differences in subjective organisation in the euthymic patients. Group differences remained in all groups when differences in subjective organisation were corrected for. This is in line with the study of verbal memory in euthymic bipolar patients mentioned previously which reported that a semantic clustering index failed to fully explain the difference between patients and controls on the California Verbal Learning Test (Deckersbach et al., 2004b).

The fact that subjective organisation does not explain the extent of the verbal memory deficit leads back to the question of why patients show verbal memory impairment. The memory deficit may reflect a 'scar' from previous mood episodes. The limited evidence of impairment in subjective organisation combined with the relatively weaker relationship between subjective organisation and recall observed in the patient groups indicate that patients organise their recall to a very similar extent to controls, but seem to benefit less from that organisation. It is not clear from the present study why patients have been unable to profit from organisation in the same way as controls. It is not known whether the process of subjective organisation takes place primarily at encoding or retrieval, or how list items become organised over repeated trials. One possibility is that recalling two items together strengthens the contingency between them and makes future recall of the same item-pair more likely. This phenomenon has been described as retrieval-based learning (Anderson and Bjork, 1994). It is one

possibility that patients benefit less from retrieval-based learning. One finding from the present study that would be consistent with this possibility is that there was a notable increase in subjective organisation between trials 6 and 7 for the control group, but a much weaker or absent effect in the patient groups. Trials 6 and 7 are the two trials on which the list is not presented prior to recall and therefore the trials when conditions for retrieval-based learning are optimal (recall on these trials is not affected by serial-order recall effects from the recent presentation of the list, such as primacy and recency effects). It is possible that the increase in organisation seen in the control group reflects a process underlying retrieval-based learning. It is necessary to explore this further in patients with bipolar disorder and investigate whether the deficit relates to establishing sufficient cue-strength to recall items when needed.

The possibility that subjective organisation is not the factor of interest in understanding memory impairment in patients with bipolar disorder is also a possible conclusion. Although some differences exist – more prominently under standard administration conditions – they are not of sufficient magnitude to explain the extent of the memory deficit in its entirety. There were several differences between the three sets of samples employed, not merely in mood state and administration condition, but in terms of when they were recruited, the person who tested them, and the version of the task they received (the depressed sample, recruited as part of a longitudinal investigation with multiple testing points, were randomised to receive one of three different versions of the list-learning task). It was not possible to eradicate all of these differences or to control for them statistically in any meaningful way, but if they contributed in some way to the variation between samples in the pattern of the findings, at the very least this would suggest that any effect of subjective organisation is sensitive and not especially robust.

The major question that emerges is what do these indices of subjective organisation actually measure and do they measure it reliably? The investigations mentioned previously in patients with frontal lobe damage provided some indication that subjective organisation is a 'frontal function' – patients with frontal damage show impairment in list-learning tasks and also lower levels of subjective organisation (Gershberg and Shimamura, 1995, Stuss et al., 1994). However, this is not definitive evidence that they capture executive processes. Throughout the analyses, the two different measures (ARC' and SO) told a subtly different story (despite showing good convergent validity). The SO measures tended to show smaller differences between patients and controls and a weaker relationship with recall (although this was not borne out in terms of correlations that were statistically significantly weaker between SO and recall than those between ARC' and recall). Part of this difference relates to their mathematical properties – ARC' corrects for chance recall of items, whereas SO is the measure that shows greatest independence from total recall (Murphy and Puff, 1982). This relates to a major difficulty with indices of subjective organisation – they are all derived from recall data, there is no independent or direct way to measure the construct otherwise and therefore no way to ascertain which index is the most valid measure. Despite deriving indices that mathematically and conceptually have the maximum possible separation from recall itself, it may well be that the degree of separation remains insufficient to truly tease apart the executive and non-executive memory processes or that enforcing this separation removes relevant and informative relationships.

There is limited data investigating whether subjective organisation indices do measure executive processes. One study in patients with Alzheimer's disease reported that lower subjective organisation in the prodromal phase (people with mild cognitive impairment) was associated with a higher risk of subsequently developing Alzheimer's

disease (Ramakers et al., 2010). However, the authors reported that the effects of subjective organisation were not mediated by any of the executive measures also included in their test battery. A study in adolescents reported that indices of serial ordering or semantic clustering on the California Verbal Learning Test correlated poorly or not at all with measures of executive function derived from factor analysis (Beebe et al., 2000). However, measures of recall correlated with the problem solving/mental flexibility factor. This suggests that although verbal recall shows a relationship with executive function, measures of organisation in output do not capture the executive component of the tasks. Exploratory analyses in the present investigation showed that in bipolar patients the subjective organisation measures correlated significantly with various other measures in the test battery, including executive and psychomotor measures. However, the pattern of relationships was not consistent and the signs of the relationships varied indicating both a positive and a negative relationship between subjective organisation and executive function. The control sample showed very limited evidence of a relationship between subjective organisation and other cognitive measures. In the patient sample, total verbal recall on the list-learning test did show a significant relationship with several other executive and psychomotor processing measures. However, some of these relationships most likely related to commonalities between the tests (e.g. verbal fluency and verbal recall may be associated by a general verbal ability factor; the significant relationship with the planning task may reflect the memory component of this particular planning task (the self-ordered pointing test)). These results are exploratory and would benefit from more detailed investigation using factor analysis or a statistical approach better-designed to identify patterns of common variance between tasks. However, taken together with these caveats in mind, the present findings raise questions about what subjective organisation indices actually measure and whether or not they reflect an executive component of learning and

memory. Potentially the hypothesis as initially stated (that memory deficits would be explained by deficits in executive function) remains untested rather than unsupported.

The issue of the reliability of subjective organisation measures is highlighted by the large degree of variation that was seen in these measures between the three control samples. There is at least as much variation between the control samples as there is between the patient and control samples. Although some of this variability may have arisen due to differences in the samples themselves, at the very least this suggests the subjective organisation indices are highly sensitive to small differences between samples and raises concerns about their reliability.

One alternative perspective on the current findings is that the importance of list-ordering was demonstrated, but not via the measurement of subjective organisation. The better test of whether a strategy is useful is seeing what happens to performance when the strategy cannot be used. If the consistent presentation order plays a key role in performance, then removing it by shuffling the list would impair performance. It was demonstrated in the study involving euthymic patients with the shuffled list presentation that the performance of control participants was relatively more affected than that of the patients, especially in the later trials. The control participants' performance continued to exceed that of the patients on each trial, but the gap had narrowed compared to that seen in the standard administration, perhaps because control participants were not able to benefit from the repeated presentation of the items in the same order. It may be that the patients benefit less from the consistent presentation in the standard administration and therefore are relatively less affected when strategies that it enables are no longer available. In essence, the patients do not miss a strategy that they are not already using. Replication and further exploration of the role that list presentation order plays in verbal recall impairment in bipolar disorder is necessary to

understand these issues. Although the subjective organisation indices did not relate to recall in a way that explained the performance deficit statistically, perhaps the behavioural data alone provide stronger evidence that bipolar patients benefit less from a consistent list-presentation order, but subjective organisation indices do not measure the relevant processes effectively.

There are a number of limitations of the present study necessary to discuss. One potential weakness is the cross-sectional nature of the study. Each of the samples was recruited separately and tested independently of each other under subtly varying conditions. The fact that the different list-presentation conditions were not tested within the same subjects leaves the possibility that differences in the pattern of findings between studies stems from inherent differences between the groups rather than the manipulations of interest. However, testing the different conditions within participants would have suffered the confound of practice effects, and matching each patient group with their own control group (in terms of age, gender and education) controlled for demographic differences on a study-by-study basis therefore minimising within-study artefacts.

The sample sizes varied across the studies, with inevitable differences in statistical power that result. Focussing on solely statistically significant differences would therefore be misleading owing to the likely higher rate of type II error in the studies with smaller samples. Including an effect size measure (Cohen's d) enabled a comparison of the scale of the measured group differences across the different samples. However, Cohen's d is a less reliable measure in small samples and it would have been ideal to have the samples a more similar size.

A further limitation is that the items on the word lists were presented at the rate of one per second, which may be too fast to promote the ideal conditions for subjective

organisation to occur (Murphy and Puff, 1982). A presentation rate of one item every two seconds spreads the list items out over a greater time span which permits more rehearsal and reduces the benefit to be had from relying solely on the immediate verbal memory trace (Murphy and Puff, 1982). This may encourage participants to invoke other strategies, such as subjective organisation, to assist their recall. It could be that presenting the words at a slower rate would reveal greater differences in subjective organisation. On a related procedural note, the group receiving the shuffled list all received the same random order of the lists. To control for any artefacts that this may have created (i.e. the fact that some list-pairs had more chance intertrial repetitions than others) different sequences of lists should have been used. However, given that patient/control group differences were the focus of interest and each group received the same presentation order, it seems unlikely that this will have had a major impact on the current findings.

In conclusion, the present study extended previous work investigating the relationship between executive function and other cognitive processes. The overall finding was that depressed and euthymic patients with bipolar disorder show impairment in verbal recall that is not wholly accounted for by differences in the subjective organisation of recall output. Questions were raised about the processes captured by subjective organisation indices, and there was no evidence that they measured executive processes. One potential hypothesis worthy of further exploration is whether patients with bipolar disorder show difficulties with retrieval-based learning. Assessing recall of material multiple times after only a single presentation, or using standard list-learning tasks and adding additional recall trials without presenting the list beforehand could begin to address this question.

CHAPTER 7: MODIFYING COGNITIVE IMPAIRMENT IN BIPOLAR DISORDER

INTRODUCTION

The current chapter is a pilot study to explore whether executive function deficits in patients with bipolar disorder can be modified. As has been discussed previously, the reasons behind cognitive impairment in bipolar disorder are not known and in euthymic patients with bipolar disorder there have been very few attempts to investigate whether cognitive dysfunction can be improved (Deckersbach et al., 2009, Khan et al., 2004).

In recent years, the notion of cognitive remediation has been a topic of great interest in patients with schizophrenia. There is evidence of substantial cognitive impairment shown by patients with Schizophrenia (Fioravanti et al., 2005, Heinrichs and Zakzanis, 1998) and an association with poor functional outcome (Green, 1996). A number of different approaches have been used to explore whether cognitive deficits in patients with Schizophrenia are amenable to intervention. For example some have used a one-off experimental modification (such as altered instructions or extrinsic motivation) to see if short term gains on a particular cognitive test can be made (Bellack et al., 1990, Goldman et al., 1992b, Green et al., 1990). Others have used an extended programme of cognitive remediation which takes place over a number of sessions with the aim of making gains in cognitive function that generalise beyond the paradigms used during the remediation programme (for a review see McGurk et al., 2007). One of the primary aims

of cognitive remediation programmes for patients with schizophrenia has been to improve functional outcome by reducing cognitive impairment. A recent meta-analysis indicated a small effect of cognitive remediation therapy on cognitive deficits in patients with schizophrenia (d=0.41) and a small effect for social function (d=0.36) (McGurk et al., 2007).

In patients with bipolar disorder, the potential role of pharmacotherapy in managing cognitive impairment in bipolar disorder has received more attention than non-pharmacological interventions (Burdick et al., 2007). To date there has only been a single study investigating a non-pharmacological intervention designed to modify cognitive dysfunction in bipolar patients (Deckersbach et al., 2009). This study investigated a 14-session cognitive remediation programme in an open trial. The programme had the dual aims of reducing inter-episode symptoms and improving cognitive function in euthymic patients with bipolar disorder. The focus was cognitive behavioural in nature, and the sessions aimed to increase mood monitoring, activity scheduling, staying 'on task', and using internal reminders. Cognitive function was assessed at baseline, but was not tested formally at follow-up (which took place three months after the end of the three month block of treatment sessions). Significant decreases in depressive symptoms, work dysfunction, and executive dysfunction (as measured by the executive subscale of the Frontal Systems Behavior Rating Scale, a behavioural measure of executive dysfunctions associated with frontal lobe pathology) were evident at the end of the three month treatment block and maintained at follow-up.

Although the results of this intervention are promising, much more needs to be known about which interventions are most effective for patients with bipolar disorder and what interventions may be doing that is beneficial to patients.

Of the experimental interventions used for patients in Schizophrenia, some have involved enhanced instructions or even telling participants how to perform the task (e.g. Green et al., 1990). The task most-commonly used for experimental interventions of this nature has been the Wisconsin Card Sorting Test (WCST). A valid criticism levelled at this approach is that if the aim of an intervention is to identify whether patients can perform the task under optimal conditions, the fundamental nature of the task needs to remain unchanged (Goldberg and Weinberger, 1994). Part of the standard administration of the WCST involves withholding the fundamental principle of the task and participants use deduction to work out the underlying rules. Telling participants the rules changes the task demands markedly. One intervention that does not suffer this limitation to the same degree is self-monitoring, that is asking a participant to say aloud their reasons for performing the task as they are. It has been used successfully in patients with Schizophrenia to improve performance on the WCST (Perry et al., 2001) and it has been used to mixed-effect on other cognitive tasks (Harvey et al., 2009). The intervention seems to work best for more complex tasks but to impair performance on simpler tasks (Harvey et al., 2009). Depending on the task used, it may be possible to identify the aspects of processing that self-monitoring improves by identifying the performance indices most enhanced by the intervention.

The WCST lends itself to use in a self-monitoring context. It requires complex processing, involving concept formation, response inhibition, set-shifting, and memory. It is a self-paced test with minimal instruction, and a diverse range of outcome measures. Performance on the WCST is impaired in patients with bipolar disorder (Robinson et al., 2006) and it has been the focus of a lot of cognitive remediation work in patients with Schizophrenia (Goldberg and Weinberger, 1994).

The present study uses a self-monitoring intervention on the WCST in euthymic patients with bipolar disorder to identify whether it may have potential as a remediating intervention in this patient group. Contrasting performance on the test by a group of patients who receive the standard administration versus those who receive a modified version requiring self-monitoring (compared both directly and against control groups receiving the same two administration conditions) will identify whether self-monitoring is associated with better performance. It is hypothesized that self-monitoring will be associated with a smaller performance decrement in Wisconsin card sorting test performance in euthymic patients with bipolar disorder than euthymic patients who receive standard administration.

METHODS

PARTICIPANTS

The details of the clinical sample are the same as those described previously in Chapter 5 on page 97.

MATERIALS

1. WISCONSIN CARD SORTING TEST – 128 CARD VERSION (HEATON ET AL., 1993)

Administered on the computer, the participant is presented with four base cards at the top of the screen – one showing a red triangle, one showing two green stars, one showing three yellow crosses, and one showing four blue circles – and a single card at the bottom of the screen which matches with at least one of the four base cards on at least one possible dimension (colour, shape, or number). Using the touch screen monitor the participant is asked to sort the bottom card to one of the base cards without being told the sorting principle (the bottom card is moved to a row beneath the base cards and

the top four cards in each 'pile' remain visible throughout the task). After each sort the word 'Right' or 'Wrong' appears on the screen depending on whether the card was correctly sorted and the participant must use trial and error to ascertain the underlying rule. Three sorting rules are used – colour, form, or number. After ten consecutive correct responses to the first category – and unknown to the participant – the rule changes and the participant must switch from the established sorting principle and again use trial and error to ascertain the new rule. After ten consecutive correct responses to the new category, the rule switches again and so on until either five shifts in category are made successfully or until all 128 cards have been dealt.

There are a number of outcome measures derived from this task: the number of trials administered, the number of correct responses, the number of trials until the first category is achieved (i.e. ten consecutive correct responses), total errors, total perseverative responses, total perseverative errors, total nonperseverative errors, the number of categories achieved, the number of conceptual level responses, failure to maintain set (the number of times after five consecutive correct sorts to a category that the participant fails to achieve the category), and an index measuring learning to learn (the average decrease in errors made with each successive category).

For a cross-sectional exploration of the impact of self-monitoring on Wisconsin Card Sorting Test performance, after 21 patients and 20 control participants had been tested using the standard instructions, the instructions were modified to ask the participant after each card sort "Why did you put that there?". The aim was to encourage the participant to 'think aloud' and to reason through their choices. The question was asked in a friendly manner and after the first few times was generally contracted as appropriate on an individual-basis depending on the response to the questioning (e.g. after a few sorts some participants provided their reasons without

prompting, whereas others required prompting throughout). The participant was not given any further prompting or any feedback about whether they had sorted for the correct reason.

PROCEDURE

The Wisconsin Card Sorting Test was administered as part of a larger battery, the details of which are provided in Chapter 5. The first twenty one patients and the first twenty controls were administered the standard version. Subsequently all patients (n=16) and controls (n=8) were administered the modified version.

DATA ANALYSIS

Data were analysed using SPSS version 17.0. Group differences were examined initially using factorial ANOVA with 2 between-groups factors each with two levels (Group (patient or control) and Condition (standard or self-monitoring)). Then, as this was a pilot study with a relatively small sample size and low numbers in some of the groups, this analysis was followed up with a less-conservative approach. The four groups (control standard, control self-monitoring, bipolar standard, bipolar self-monitoring) were compared using one-way ANOVA. Differences reaching statistical significance at $p \leq 0.05$ were followed up with post-hoc LSD tests, again using $p \leq 0.05$ as the criterion for statistical significance. As in the previous chapters, variables were screened for outliers and any showing extreme outliers were compared using a non-parametric equivalent to one-way ANOVA, the Kruskal-Wallis test.

Effect sizes were calculated using Cohen's d where d=(control mean-patient mean)/pooled standard deviation. Effect sizes for indices where a higher score indicated worse performance (e.g. error or reaction time indices) were reflected (multiplied by -1) such that a positive effect size always indicates a poorer performance by the patient

group. Conventions for Cohen's d are that 0.2<d<0.5 is a small effect, 0.5≤d<0.8 is a medium effect, and d≥0.8 is a large effect.

RESULTS

Performance of the four groups is shown in Table 7.31 below.

The results of the 2x2 factorial ANOVA indicated a statistically significant main effect of Group on six of the eleven outcome measures, all indicating poorer performance by the patient group (trials administered, $F_{1,61}$=6.81, p=0.011; total errors, $F_{1,61}$=5.38, p=0.024; nonperseverative errors, $F_{1,61}$=6.82, p=0.011; categories achieved, $F_{1,61}$=6.57, p=0.013; trials to 1st category, $F_{1,61}$=5.18, p=0.026; failure to maintain set, $F_{1,61}$=8.78, p=0.004). Two further indices showed a trend towards a statistically significant main effect of Group, again reflecting poorer performance by the patients (perseverative responses, $F_{1,61}$=2.87, p=0.095; perseverative errors, $F_{1,61}$=3.51, 0.066).

There were no main effects of Condition for any index (all $F_{1,61}$<2.3, all p>0.14).

There was one Group x Condition interaction that showed a trend towards statistical significance (total trials correct, $F_{1,61}$=2.92, p=0.093), but none of the others approached statistical significance (all $F_{1,61}$<2.6, all p>0.11).

Given the small sample size, following these broadly negative results up with a less conservative analysis (one-way ANOVA) revealed significant differences between the groups on several outcome measures, including the number of trials administered (F=4.50, p=0.006), non-perseverative errors (F=3.24, p=0.028), categories achieved (F=3.02, p=0.036), trials to first category (F=4.03, p=0.011), and failure to maintain set (F=5.81, p=0.001).

Table 7.31: Performance on the Wisconsin Card Sorting Test. Means and standard deviations of the four groups. The F and p values relate to the results of the one-way ANOVA analysis.

	Control standard (n=20)		Bipolar standard (n=21)		Bipolar self-monitoring (n=16)		Control self-monitoring (n=8)			
	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	F	p
Trials administered	88.75	19.10	112.05	18.70	100.88	23.61	95.38	23.18	4.50	0.006
Total trials correct	66.20	8.52	70.81	12.66	67.19	11.32	72.50	10.11	1.04	0.380
Total errors	22.55	21.92	41.24	24.64	32.75	25.75	22.88	13.37	2.63	0.058
Perseverative responses	13.70	16.96	21.52	14.27	19.94	19.23	13.25	9.29	1.12	0.347
Perseverative errors	11.80	12.87	19.67	12.47	17.06	15.74	11.88	7.77	1.53	0.216
Nonperseverative errors	10.75	11.63	21.57	13.26	17.00	13.09	11.00	5.71	3.24	0.028
Conceptual level responses	61.40	12.72	57.81	20.77	56.81	17.45	66.13	4.02	0.74	0.532
Categories achieved	5.50	1.40	4.00	2.32	4.75	1.95	5.75	0.46	3.02	0.036
Trials to 1st category	12.80	3.00	35.81	37.04	17.75	20.78	12.00	2.45	4.03	0.011
Failure to maintain set	0.20	0.41	1.24	1.30	0.56	0.63	0.25	0.46	5.81	0.001
Learning to learn	3.71	26.30	25.07	49.73	2.18	30.87	-0.86	3.25	1.98	0.127

Table 7.32 shows the results of the post-hoc tests for the three key comparison-pairs of interest (controls with standard administration versus patients with standard administration; controls with standard administration versus patients with modified administration; patients with standard administration versus patients with modified administration; NB neither patient group was contrasted with the control group who received modified administration owing to the small sample size in the latter group restricting statistical power). Patients who received standard administration showed significantly poorer performance than controls on all five measures (all $p<0.05$), whereas patients who received the self-monitoring intervention showed no statistically significant differences in performance from the control group. Furthermore, on two of the measures (trials to first category and failure to maintain set) patients who engaged in

self-monitoring showed significantly better performance than patients who had the standard administration.

Table 7.32: Post-hoc tests following up the one-way ANOVA, p-values of the LSD test

	Ctrl std vs BD std	Ctrl std vs BD sm	BD std vs BD sm
Trials administered	0.001	0.085	0.108
Nonperseverative errors	0.006	0.128	0.258
Categories achieved	0.011	0.226	0.221
Trials to 1st category	0.003	0.535	0.025
Failure to maintain set	0.000	0.210	0.020

Ctrl, control group; BD, bipolar disorder group; std, standard administration; sm, self-monitoring

The effect sizes for the three key between-group comparisons are reported in Table 7.33 below. In general, the patients who received the standard administration showed medium to large impairment on the task indices, whereas the effect sizes for patients who used self-monitoring were small to medium (in all but one instance indicating poorer performance than controls). The two patient groups when contrasted with each other showed small to medium effects, in all but one instance indicating better performance by the self-monitoring group.

Table 7.33: Effect sizes (Cohen's d) on the Wisconsin Cart Sorting Test

	Ctrl std vs BD std	Ctrl std vs BD sm	BD std vs BD sm[a]
Trials administered *	1.23	0.57	0.53
Total trials correct	0.43	0.10	0.30
Total errors	0.80	0.43	0.34
Perseverative responses	0.50	0.35	0.10
Perseverative errors	0.62	0.37	0.19
Nonperseverative errors *	0.87	0.51	0.35
Conceptual level responses	0.21	0.31	-0.05
Categories achieved *	0.78	0.45	0.35
Trials to 1st category *	0.86	0.35	0.58
Failure to maintain set *	1.07	0.70	0.63
Learning to learn	0.53	-0.05	0.54

[a] A positive effect size indicates poorer performance by the standard administration group
* indicates omnibus F-test was significant at $p<0.05$ level
Ctrl, control group; BD, bipolar disorder group; std, standard administration; sm, self-monitoring

To establish whether these differences were related to demographic differences or differences in cognitive ability between the four groups, further ANOVA analyses were undertaken comparing the groups on age, years of education, premorbid IQ, and a number of executive measures (phonetic fluency, category fluency, Hayling sentence completion test, Zoo map, reverse digit span, trail making test part B, Stroop, and the self-ordered pointing task). There were no significant differences between the two patient groups on any of these measures (all $p>0.05$; data not shown), except the zoo map test – patients who received the standard administration were significantly worse than patients who received the self-monitoring intervention (Zoo map version 1 raw score, $p=0.006$). Comparing the two patient groups on the five Wisconsin Card Sorting Tests on which they showed significant differences (see Table 7.32) using ANCOVA to

control for Zoo map score removed all significant differences on the WCST measures (all p>0.126).

DISCUSSION

The results indicate that self-monitoring is associated with a smaller deficit on the WCST in euthymic patients with bipolar disorder. Patients who received standard administration showed significant impairment compared to controls on five out of eleven test indices. However, those who received the self-monitoring intervention showed no significant differences from controls. On two indices, the patients who had engaged in self-monitoring significantly out-performed the patients who completed the standard administration.

This is a novel finding in this patient group. A relatively straightforward intervention that could be easily encouraged or used in other situations was associated with reduced cognitive impairment.

The question turns to the mechanism by which self-monitoring may assist processing. Previous authors have commented that it may exert its effects via reducing distractibility and increasing focus on the task at hand (Perry et al., 2001). There is some suggestion from the present study that it aided performance by supporting working memory (e.g. helping participants remember which category they were sorting by) and by assisting concept formation. When contrasting the two patient groups, the areas where the largest effect sizes were noted was in failure to maintain set (potentially indicative of simply 'forgetting' which category was the sorting principle) and trials to first category (potentially reflecting the trial and error process used to derive the concepts underlying the task). Another way of thinking about the mechanism by which it may be supporting performance is that self-monitoring may create a 'social

accountability' for the decisions made, enforcing reasoning to become conscious and deliberate. It is difficult to test this conjecture, as it impossible to know the reasoning processes underway in the absence of verbalisation. However, it is possible that by requiring verbalisation, thoughts which may otherwise remain unstructured are forced (by virtue of needing to understood by another) into some form of structure, which in itself is helpful for ongoing processing. In the developmental psychology literature, the social learning theorist Lev Vygotsky observed that young children frequently talk aloud to themselves and conjectured that this so-called ego-centric speech is used functionally for problem solving (Vygotsky, 1962). Observational studies of children showed that the rate of ego-centric speech almost doubles when they encounter a difficulty in a task, indicating that ego-centric speech "becomes an instrument of thought in the proper sense – in seeking and planning the solution of a problem." (p. 16, Vygotsky, 1962) By encouraging the reactivation of this strategy, it seems benefits may also be had for adults. Interestingly, however, the self-monitoring intervention was not associated with improved performance in the control participants. Although the sample size was too small to draw any definitive conclusions, the present results indicated that in those with 'normal' performance further improvements are not made by encouraging self-monitoring.

Do the current findings have any implications for understanding the nature of the underlying deficit in bipolar disorder? It is tempting to conclude, as has been done for patients with schizophrenia, that the impact of psychological interventions mitigates against the likelihood that the deficit is due to immutable structural (or functional) damage or a fundamentally biological cause (Summerfelt et al., 1991). However, caution is required in accepting this possibility. Without developing a detailed understanding of how the intervention is helping, it is not possible to draw any conclusions about the nature of the underlying deficit. Authors have suggested that the fact performance on a

tests such as the WCST can be improved using short-term experimental manipulations may stem in part from the fact that all cognitive measures are an imperfect estimate of the underlying functions (Goldberg and Weinberger, 1994). They are not only subject to measurement error, but may also suffer "functional redundancy" (Goldberg and Weinberger, 1994, p. 294). That is, a (statistically) normal result may be possible even if the underlying cognitive systems used by the task are damaged as the task does not require the underlying circuitry to be used to its full capacity. The key message from the current study is that cognitive dysfunction in euthymic bipolar patients shows a degree of modifiability. This suggests that extended programmes designed to develop the use and internalisation of self-monitoring processes by bipolar patients without external prompting may be useful for ameliorating cognitive impairment in this group and should be explored further.

It is also necessary to bear in mind that it is not the case that this simple intervention normalised performance on the task. Firstly, the cross-sectional nature of the data does not permit this conclusion. This was not a within-subjects design, and although the between-subjects design avoided the contamination from practice effects (or effects from establishing a negative response set), it remains impossible to determine whether performance changed from impaired to unimpaired in the patients who received the self-monitoring intervention. Secondly, in general the intervention group showed no significant differences from either the control group or the patients who received standard administration (the aforementioned indices excepted). This suggests their performance was intermediate between the two groups rather than normalised to the point that there was a significant performance advantage over and above patients who received standard administration. Also, the effect sizes observed in the patients who received the modified administration contrasted with controls who received the standard administration still remained in the small to moderate range. In a larger

sample, these differences would have reached statistical significance, and would result in a deficit of a similar magnitude to that generally found for verbal fluency tasks or short term memory measures in euthymic patients with bipolar disorder.

Some of the limitations of the current findings have already been mentioned, most notably the cross-sectional design. Additionally patients were not randomised to standard administration or self-monitoring; the groups were tested serially. Fundamental differences in the groups may thus have accounted for part of the findings rather than merely the intervention itself. Indeed preliminary analyses investigating whether there were differences on other cognitive measures between the patient groups identified that the patients who received the standard administration were worse at a planning task than the patients who received the self-monitoring intervention. Controlling for differences on this task meant that differences on the WCST were no longer statistically significant between the groups. It is therefore necessary to replicate these findings in a larger sample using a randomised design. The small sample size in some groups (especially the controls who received modified administration) undoubtedly affected statistical power and replication in a larger sample is necessary. The initial findings of the factorial ANOVA indicated broadly negative results, likely due to poor statistical power. Therefore this should be considered a promising pilot study until the results can be replicated in a larger sample size with an improved study design. Finally, recording the responses of the participants would have provided valuable information for understanding how participants reasoned through the task. In general the intervention was well-received, but the repetitive nature of the questioning was a source of frustration for some participants. There may be more effective ways of encouraging verbalisation which avoid this difficulty.

Despite these limitations, there is nonetheless evidence that a straightforward and simple intervention was associated with a smaller performance decrement on WCST. This is a very optimistic message and further research should clarify to what extent it can be useful for patients in tackling difficulties arising from cognitive dysfunction in everyday life.

CHAPTER 8: EMOTION PROCESSING IN BIPOLAR DISORDER

INTRODUCTION

The focus of the current chapter is emotion processing in bipolar disorder. As discussed in Chapter 3, it is an area that is relatively underexplored in contrast with neuropsychological function in this group. Given that mood disturbance is one of the defining features of the illness, and that emotions play a vital role in interpersonal functioning, it is a key area of investigation.

FACIAL EXPRESSION RECOGNITION

Chapter 3 provided a review of the evidence of emotion processing deficits in individuals with bipolar disorder. As a brief recap, previous studies have noted differences in the response pattern of patients with bipolar disorder in facial emotion labelling paradigms in manic, depressed and euthymic patients (Getz et al., 2003, Gray et al., 2006, Harmer et al., 2002, Lembke and Ketter, 2002). The pattern has not been consistent across studies. Studies in euthymic patients have identified enhanced recognition of fear (Lembke and Ketter, 2002), enhanced recognition of disgust (Harmer et al., 2002), impaired recognition of fear (Venn et al., 2004), and no deficit in emotion recognition (Vaskinn et al., 2007). However, these studies have generally not been direct replications of one another, involving a variety of different methodologies and differences in the characteristics of the patient samples investigated.

The present study uses a facial expression recognition task using still facial images of individuals displaying one of five different facial expressions (angry, disgusted, fearful, happy, and sad) presented at four different intensities of expression (20%, 40%, 60%, 80%). Neutral faces (and a neutral response option) are also included. Still facial images were chosen to avoid artefacts caused by the non-naturalistic way expressions are formed in the tasks involving moving images. The specific intensities were selected to sample more highly from images where the expression is more ambiguous in order to best-expose potential biases in interpretation.

As a test of recognition of emotions other than the basic emotions, the Reading the Mind in the Eyes Test (mentioned previously in Chapter 3) is also included.

VOCAL EMOTION RECOGNITION

Irrespective of whether patients show deficits in facial emotion recognition, it remains interesting to explore whether recognition of emotion via other modalities is altered in bipolar disorder. The present study includes a vocal emotion recognition paradigm that assesses recognition of anger, disgust, fear, happiness, surprise and sadness from vocal stimuli.

IMPLICIT EMOTION PROCESSING

The present study employs an emotional Stroop paradigm and a modified dot probe task to further extend previous studies of using similar measures in patients with bipolar disorder.

Previous studies using the emotional Stroop paradigm in manic patients with bipolar disorder showed greater interference for negative or depression-related words than positive or mania-related words. It is unclear whether a similar bias exists in

euthymic patients (the only study using the task in euthymic patients did not explore interference effects (Malhi et al., 2005)). Additionally, with regard to the Stroop paradigm, McKenna & Sharma (2004) proposed two different mechanisms that may operate in the semantic Stroop, namely a quick and a slow component (McKenna and Sharma, 2004). The quick component involves the immediate effect or interference caused by the stimulus, whereas the slow component results from repeated exposure to words from the same category gradually exerting an effect over time. The way the semantic Stroop has traditionally been administered – i.e. as a card of words with each card representing a single semantic category – confounds these two effects and the total reading time derived via this method also includes the time taken making (or, more precisely, correcting) errors. In the present study, the words are presented individually on the computer screen with the reaction time to each individual stimulus recorded separately (via a vocal response time recorder). The words from the three semantic categories are presented in random order, thus exclusively tapping the proposed 'quick' component identified by McKenna & Sharma (2004).

The current study uses a facial modified dot probe paradigm to explore attentional engagement with angry, disgusted, happy and sad facial expressions. These emotions were chosen as enhanced engagement with angry faces has been reported previously in depressed patients with bipolar disorder using a similar task (Leyman et al., 2009). Enhanced recognition of disgust has been reported in euthymic patients with bipolar disorder, which may be associated with differences in how disgusted facial expressions capture or hold attention (Harmer et al., 2002). Happy and sad expressions were included to investigate proposed mood-congruent biases. A bias towards happiness may be indicative of vulnerability to mania and a bias towards sadness may indicate vulnerability to depression.

METHODS

PARTICIPANTS

The details of the clinical sample are the same as those described previously in Chapter 5, page 97.

MATERIALS

FACIAL EMOTION EXPRESSION RECOGNITION TASK

The task used was designed for the present study based on versions used in earlier studies (Harmer et al., 2002, Montagne et al., 2007). Participants were presented with a black and white still facial photograph of a person showing one of 5 facial expressions (angry, disgusted, fearful, happy, or sad) or neutral. The images used were drawn from the Eckman series (Ekman and Friesen, 1976). Each of the expressions (apart from neutral) was shown sixteen times – four times at each of four intensity levels (20%, 40%, 60%, 80%) (5 emotions x 16 presentations = 80 stimuli). As there were not different intensities of the neutral face, the neutral expression was shown four times in total (80 + 4 = 84 stimuli in total). Four different individuals were used for the four presentations, two male and two female. The words 'Angry', 'Disgusted', 'Fearful', 'Happy', 'Sad' and 'Neutral' were presented on the screen, written on grey buttons, and presented underneath one another listed in alphabetical order on the right hand side of the screen (the buttons themselves were 160 pixels wide by 72 pixels high, however to allow for imprecision, the touch-sensitive area of screen around each button was larger measuring 180 pixels wide by 100 pixels high). The picture of the face was presented on the left hand of the screen for one second (see Figure 8.2). All of the faces were masked idiosyncratically to cover hairstyles, ears, neck etc, and presented on a black background

(333 pixels wide x 482 pixels high). After the face had displayed for one second, it was replaced with a solid black rectangle and the participant was instructed to touch the word on the screen that they thought best described the individual's expression (see Figure 8.3). It was not possible for a response to be given when the face was still being displayed. In order for participants to familiarise themselves with the position and order of the response options, the task began with six practice trials. This involved six presentations of 100% intensity of each of the 5 emotions and one neutral face. The six pictures were all of the same individual, who was not used again at any point during the remainder of the task. The practice trials were presented in a fixed order to all participants (angry, disgusted, fearful, happy, sad, neutral). The 84 experimental trials were presented in an entirely random order and each participant received a different random order.

Stimuli were presented using Superlab 4.0 (Cedrus) and responses were recorded using 15" CTX resistive touch screen LCD monitor. Responses were self-paced with the next stimulus appearing only after the participant had responded to the previous stimulus. Void responses were recorded if a participant touched the screen outside of the areas designated for each response option (i.e. touched outside of the touch-sensitive area associated with each of the buttons displaying the response options). Although this was logged as a valid response by the computer, it was not possible to classify the response especially if it occurred between two response options.

The outcome measures of interest are the number of correct responses at each intensity level for each emotion. Reaction time was not analysed as participants were not instructed to respond as quickly as possible.

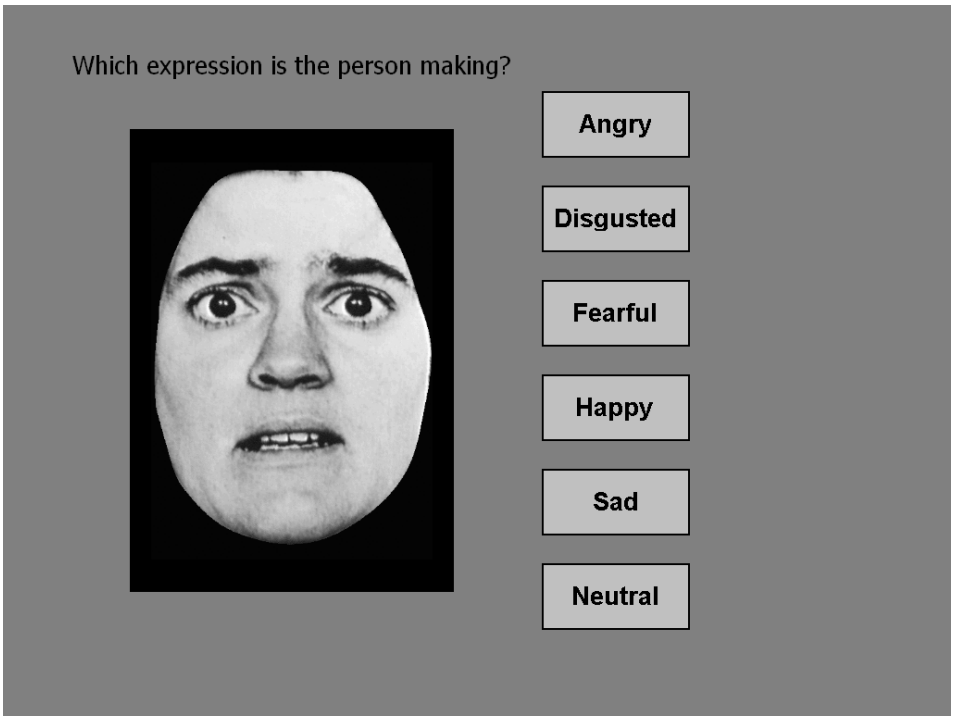


Figure 8.2: Display of the emotional face in the facial expression recognition paradigm

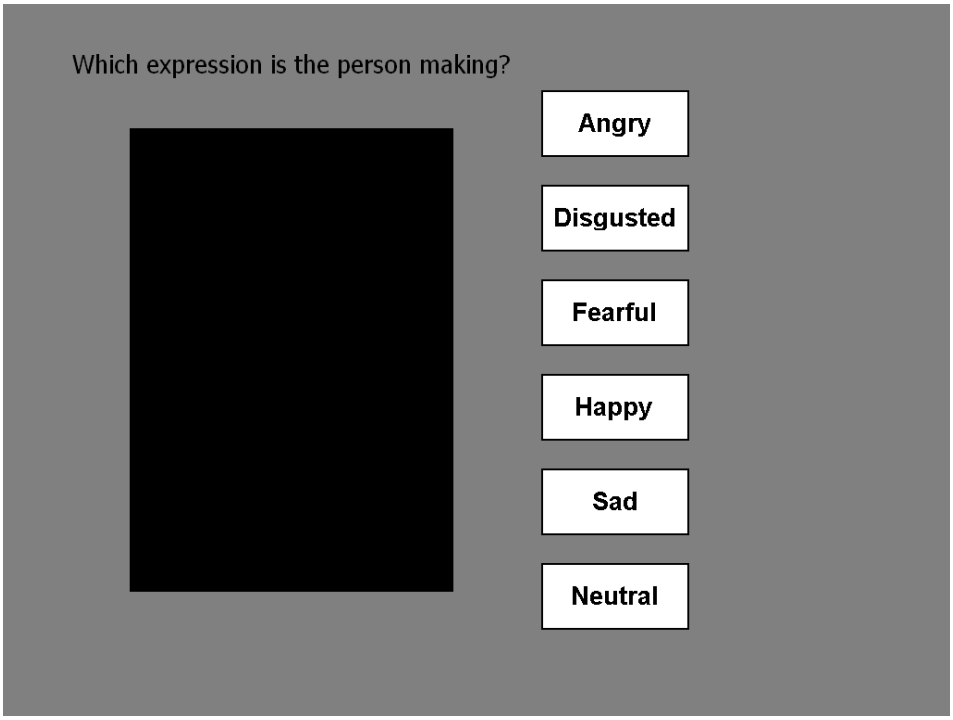


Figure 8.3: Response screen in the facial expression recognition paradigm

VOCAL EMOTION RECOGNITION TASK

This task was designed for the present study. An amateur actor was used to record six (semantically) neutral sentences in each of seven different tones of voice reflecting six emotions (angry, disgusted, fearful, happy, sad, surprised) and neutral. Participants listened to each of the sound clips while the display read 'Listen' (see Figure 8.4) and, in a similar fashion to the facial emotion recognition task, touched the emotion-word on the screen that they thought best conveyed the emotion being expressed after the clip had finished (see Figure 8.5). Seven recordings of a single sentence were used as practice, then the experimental trials (n=35) followed in a random order. Each participant received the same random order.

Stimuli were presented using Superlab 4.0 (Cedrus) and responses were recorded using 15" CTX resistive touch screen LCD monitor. Responses were self-paced with the next stimulus appearing only after the participant had responded to the previous stimulus. Void responses were logged in the same manner as for the facial expression recognition task.

The outcome measures of interest were accuracy in identifying each emotion. As the task was new and no metric was used to examine its validity, groups were additionally compared on the proportion of times each emotion-word was used to describe each sound clip. This reflects the extent to which the groups share a common interpretation of the stimuli rather than whether the answers were 'correct'.

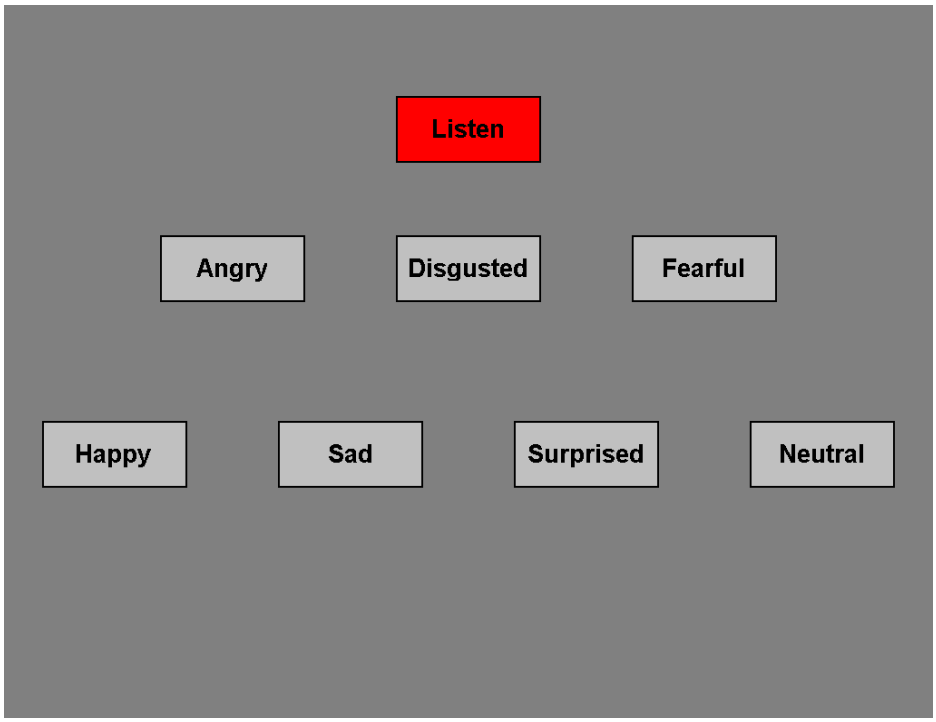


Figure 8.4: Display screen in the vocal emotion recognition task during stimulus presentation

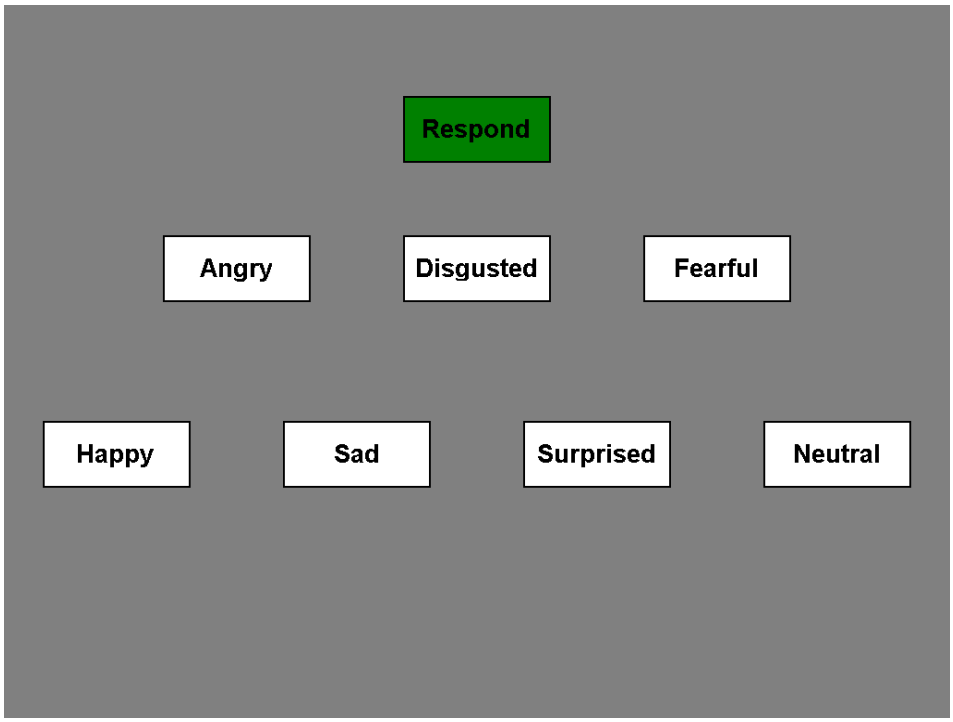


Figure 8.5: Response screen in the vocal emotion recognition task

READING THE MIND IN THE EYES TEST

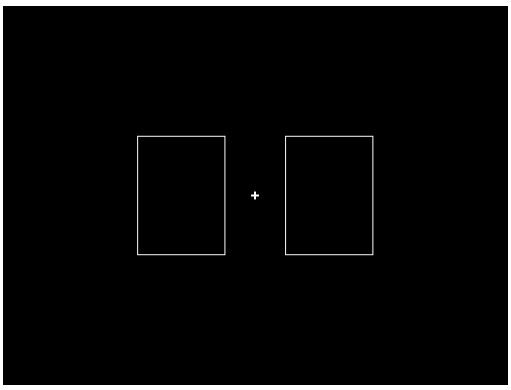
This task is described in detail by Baron-Cohen et al (2001) (Baron-Cohen et al., 2001). The participant is shown a single picture on an A4 page of the eye region of a face. The picture is surrounded by four adjectives describing a mental state (e.g. perplexed, horrified, astonished, intrigued). The participant is instructed to identify which of the words they think best describes what the person in the picture is thinking or feeling and circle their choice on a separate answer sheet. After a single practice item, 36 experimental items are completed one after the other in a self-paced manner. Response time is not recorded. The outcome measure of interest is the number of correct responses.

EMOTIONAL DOT PROBE

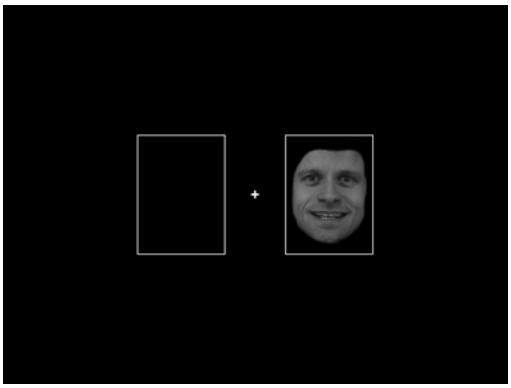
This task was designed specifically for the present study, but relates closely to similar paradigms used in previous studies (Koster et al., 2005). It is based around a Posner-style attentional cuing paradigm. Participants see a cue in one of two locations – to the left or right of a central fixation – which is then followed by a target, again in either of the same two locations. The participant responds by indicating which side the target appears. In the present task, the cue was an emotional face, and the target a small grey square. Trials consisted of two types – valid and invalid. In a valid trial, the emotional face cue and the target grey square appear in the same location one after the other, i.e. the cue validly signals the location of the target. In an invalid trial, the cue and the target appear on opposite sides, the cue is therefore an invalid signal for the subsequent location of the target square. If the cue captures attention, then reaction times to valid trials will be faster than those to invalid trials, and vice versa for a cue that fails to capture (or repels) attention. Examining the relative reaction times therefore provides clues as to which stimuli are salient to an individual.

In the current task one trial consisted of a white central fixation cross (Arial size 40) on a black background flanked by two empty white rectangles (179 pixels wide x 242 pixels high) presented for 500ms, followed by a cue face presented in either the right or the left hand box for 500ms, followed by the grey target square (10 pixels square) which remained on screen until the participant responded (see Figure 8.6). The 500ms stimulus presentation was in contrast with previous studies that used two different time intervals (200ms and 1000ms, e.g. Koster et al., 2006). The 500ms was chosen in the present study to exceed the stimulus presentation most often associated with anxiety-related processing biases (≤200ms) and to be less than the 1-second presentation time that has been suggested to permit ruminative processing (Mogg and Bradley, 2005). Faces were shown in black and white, idiosyncratically masked to hide hair, ears, neck etc., and showing one of five possible expressions – angry, disgusted, happy, sad, or neutral. The faces were drawn from the Karolinska set (Goeleven et al., 2008) to avoid any overlap with faces used in the facial expression recognition task (described above). In total, participants completed 120 experimental trials. Each expression was used 24 times – 12 times in valid trials, 12 times in invalid trials (balanced evenly for left/right presentation, left/right response). Six individuals were used for the photographs – three female, three male. Each individual was shown four times for each emotion (valid left, valid right, invalid target left, invalid target right) (6 individuals x four presentations x five expressions = 120).

a)



b)



c)

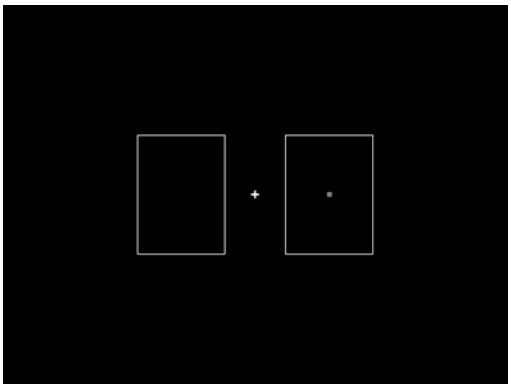


Figure 8.6: Example of a valid trial. a) presented for 500ms, b) presented for 500ms, c) presented until participant response made

Standardised instructions were given to each participant. They were presented on the screen before the task began. Following the written instructions, there was a 'walk through' trial to demonstrate the trial procedure and explain the requirements followed by six practice trials (one of each trial type on the left and the right, plus two extra). At the end of the practice trials, there was a pause and an opportunity to indicate if further practice or explanation was needed. If not, the experimental trials began when the participant pressed the designated button. Unknown to the participant the true experimental trials were prefaced with 12 'run-in' trials, used to allow the participant to 'get into their stride' before data collection began. From the participant's perspective these trials resembled any other and although data from these trials was recorded it was discarded before analysis. To avoid the task becoming a sustained attention paradigm, there was a break half way through and written instructions appeared on the screen for how to continue the task (with a button-press) when ready. The 12 trials following the break were identical to the first 12 trials and were also discarded before analysis. The faces used both in the walk through trial, practice trials, and run-in trials involved pictures of individuals who did not appear in the actual task or in the recognition task (see below). The 120 experimental trials were presented in quasi-random order, blocked into four sets of 30 trials balanced as closely as possible to include equal numbers of trials involving each of the five expressions, valid and invalid trials, right and left responses. The blocks were presented in a fixed order to each participant, but trials were randomised within each block and each participant received a different random ordering.

Stimuli were presented using Superlab 2.0.4 (Cedrus) and responses were recorded via a 6-button serial response box.

The outcome measures of interest comprise measures of attentional engagement and disengagement with each of the different emotional expressions.

Engagement contrasts the response time for valid trials displaying a neutral face with response time for valid trials showing an emotional face. In line with Koster et al 2005, it is calculated for each emotion separately using the following formula:

Engagement = average reaction time for valid neutral trials – average reaction time for valid trials of a specific emotion (i.e. angry, sad, disgusted, or happy)

Positive values indicate that participants showed enhanced engagement with the emotional stimulus relative to the neutral face, as they responded relatively faster when the target was presented in the same location as the emotional face than when the target was in the same location as the neutral face. This represents an enhanced cue validity effect for neutral faces and theoretically indicates that the emotional stimulus captured attention more effectively.

Disengagement contrasts the response time for invalid trials displaying a neutral face with response time for invalid trials showing an emotional face. In line with Koster et al (2005), it is calculated for each emotion separately using the following formula:

Disengagement = average reaction time for invalid trials of a specific emotion – average reaction time for invalid neutral trials

Positive values indicate greater difficulty disengaging from the emotional stimuli compared to the neutral faces. Invalid trials require a switch in attention from the cued location to the target location. If a stimulus is sufficiently potent to make this 'unhooking' of attention difficult, then that will be reflected in a slower response time. If the relative response time for the emotional faces is slower than for the neutral faces,

that will be reflected in a positive disengagement score. Negative values indicate faster disengagement from the emotional stimuli relative to the neutral images.

To try and prevent participants simply ignoring the face stimulus, instructions were given to "Try and pay attention to the face as much as possible as you will be asked at the end if you can recognise any of the faces you have seen." Participants were not instructed to memorise the faces or directed as to which aspects of the face to pay attention to. Immediately following the dot probe task, there was a recognition trial in which a series of 60 faces were displayed to the participants – 30 were those seen during the task (each of the six individuals showing each of the five different expressions) and 30 were faces drawn from the same set (masked and presented in the same way) but of individuals who had not been shown in the experimental trials or the practice/run-in trials or the facial emotion recognition task described above. The 30 distractor stimuli mirrored the experimental stimuli, comprising three males and three females with each of the six individuals showing each of the five expressions. Each face was shown in the centre of the screen in a white frame (of the same size used in the dot probe task) with the word 'Yes' written in white font underneath the bottom right hand corner of the frame and the word 'No' under the left hand corner (see Figure 8.7). The participant was instructed to indicate whether they think they saw the face during the preceding task and to press the rightmost button on the response box for 'Yes' or the leftmost button for 'No'. The number of hits, misses, false alarms and correct rejections were calculated for each emotion.

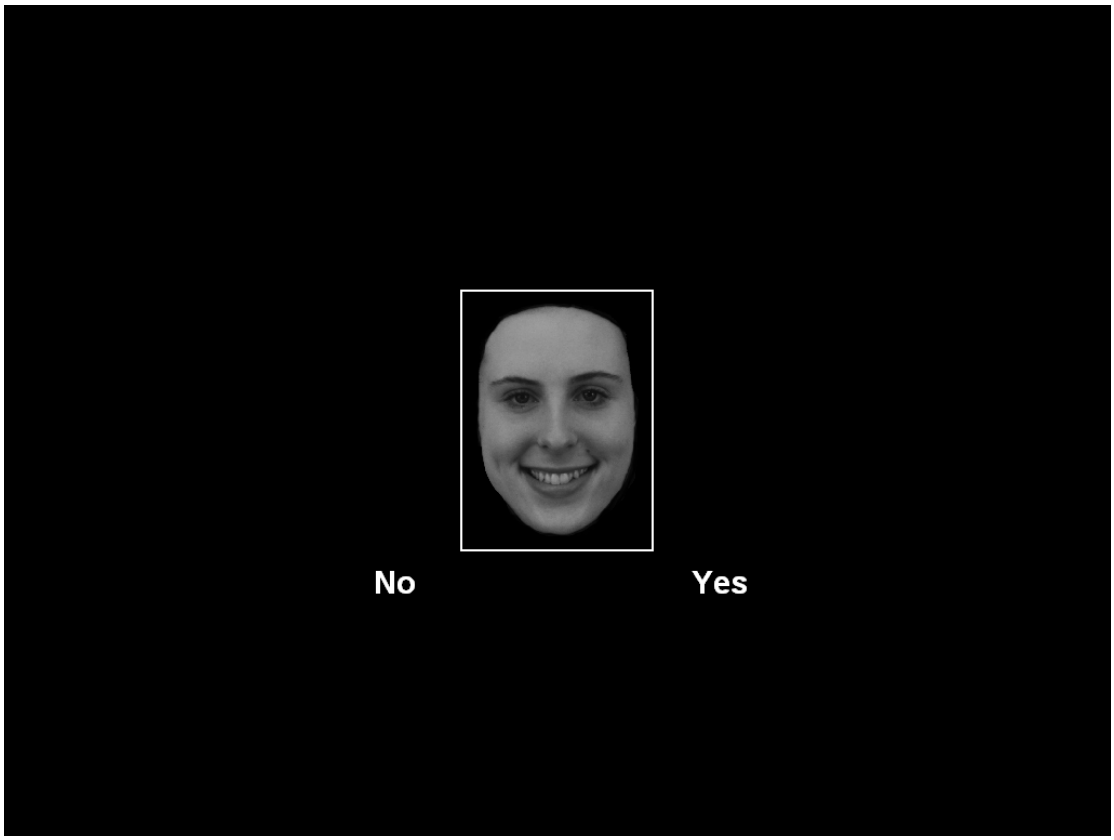


Figure 8.7: Display used in the facial dot probe recognition task

SEMANTIC STROOP

This task was presented in the same manner as the Stroop task previously described in Chapter 5, page 112. The semantic condition was presented following the administration of the 'standard' Stroop paradigm.

Participants were shown a series of 48 words belonging to one of three semantic categories – depression-related words (abandoned, afraid, agony, bad, depressed, guilty, hopeless, lonely, lost, pain, powerless, rejected, suffering, suicide, tormented, vulnerable), mania-related words (buzzing, clever, confident, creative, ecstasy, energised, euphoric, fun, gifted, happy, high, jubilant, passionate, sexy, talented, wonderful), or neutral words (amenable, assisting, blank, brave, careful, forwards, full, handy, honest, mobile, steady, talking, tall, theoretical, thoughtful, useful). The words were shown centrally in Arial size 30 in one of four colours – red, blue, green, or yellow.

Participants were instructed to say aloud the colour the word was written in. The words were presented in a quasi-random order, fixed such that each quarter of the task contained an equal number of the three categories and the four colours. Each participant received the same random order. Stimuli were presented using Superlab version 4.0 and responses were recorded via a Voice Key (Cedrus).

The three lists of words were matched overall on length, frequency of use, number of phonemes, number of syllables, and mean naming time (all $p>0.2$) (Louis, 2006). The words were rated independently by 12 raters (selected from staff and students in the Department of Psychiatry) on a scale of -7 (strongly associated with depression/low mood) to +7 (strongly associated with mania/elevated mood) with a zero point of neutral to ascertain whether the words formed distinct semantic categories. The mean ratings for each of the categories were: depression-related, -4.3 (s.d. 1.9, range of mean ratings for each word: -6.3 to -3.0); mania-related, 3.7 (s.d. 1.7, range of mean ratings for each word: +1.3 to +6.2); and neutral 0.7 (s.d. 1.5, range of mean ratings for each word: -3.4 to +2.8). There were significant differences between each of the three groups of words (all $p<0.001$).

The outcome measures comprised the average reading time for each of the three conditions, the difference in mean response time between each of the three categories, and the number of errors. Response time calculations only included correct responses. Before analyses, reaction time data was adjusted for outliers and responses <100ms or >3s.d. above the individual's own mean for any given category of words were excluded.

PROCEDURE

The tests were administered as part of a larger battery (see details in Chapter 5)

DATA ANALYSIS

Data were analysed using SPSS version 17.0. The significance level was set to $p \leq 0.05$, with $p<0.1$ representing a trend toward significance. The groups were compared using independent samples t-tests, or repeated measures ANOVA for tests that involved multiple levels or repetitions. For t-tests, Levene's F-test was first used to identify instances of unequal variance, and if $p<0.05$ corrected p-values were reported. As part of data screening, the distributions of each variable were examined using boxplots. Any variables showing evidence of extreme outliers (values more than three times the inter-quartile range) were analysed using an appropriate nonparametric test (Mann-Whitney U test). If the results of the nonparametric test differed in terms of statistical significance from the parametric test, then the former was reported. Categorical data were analysed using a Chi-squared test. Within-group comparisons (e.g. of reaction times to different categories of stimuli) were tested with paired t-tests.

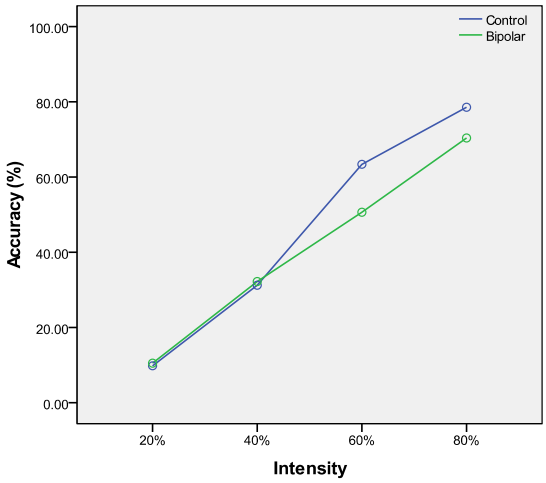
RESULTS

FACIAL EXPRESSION RECOGNITION TASK

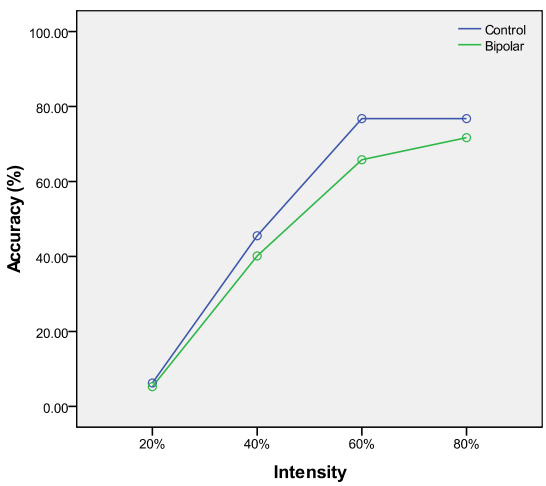
The results of the facial expression recognition task are shown in Figure 8.9 and in Table 8.34 below. The results of the repeated measures ANOVAs (run for each emotion separately with one between subjects factor (group, 2 levels) and one within subjects factor (intensity, 4 levels)) indicated that there were no significant differences between the groups for any of the five emotions (main effect of group for anger, $F_{1,64}=1.69$, $p=0.20$; disgust, $F_{1,64}=2.11$, $p=0.15$; fear, $F_{1,64}=0.25$, $p=0.62$; happy, $F_{1,64}=1.11$, $p=0.30$; sad, $F_{1,64}=0.06$, $p=0.81$). Using an independent samples t-test, there were no significant differences between the two groups for recognition of neutral faces ($t_{64}=0.81$,

p=0.42). There were main effects of intensity for each of the five emotions (all $F_{3,192}$>68.99, all p<0.001), but there were no significant group by intensity interaction effects for any emotion (anger, $F_{3,192}$=2.04, p=0.11; disgust, $F_{3,192}$=0.69, p=0.56; fear, $F_{3,192}$=1.31, p=0.27; happy, $F_{3,192}$=1.75, p=0.16; sad, $F_{3,192}$=0.05, p=0.99).

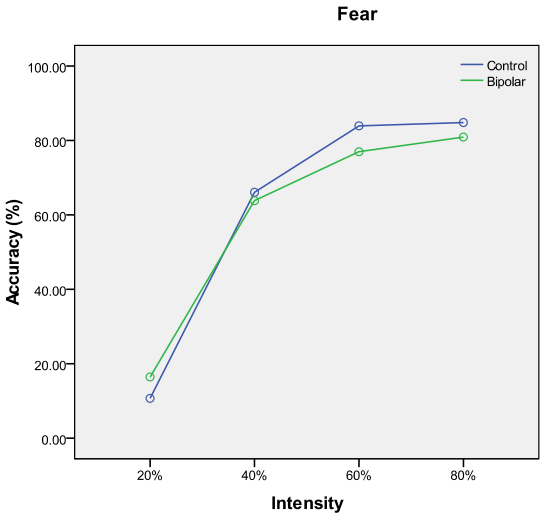
a)



b)



c)



d)

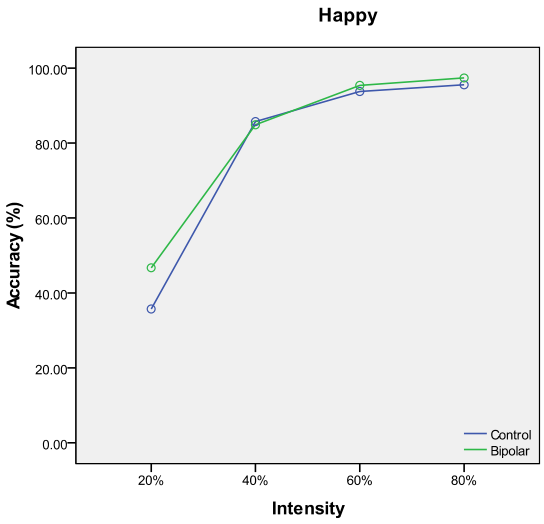


Figure 8.8a-d: Performance on the facial expression recognition test. Mean % accuracy at each level

e)

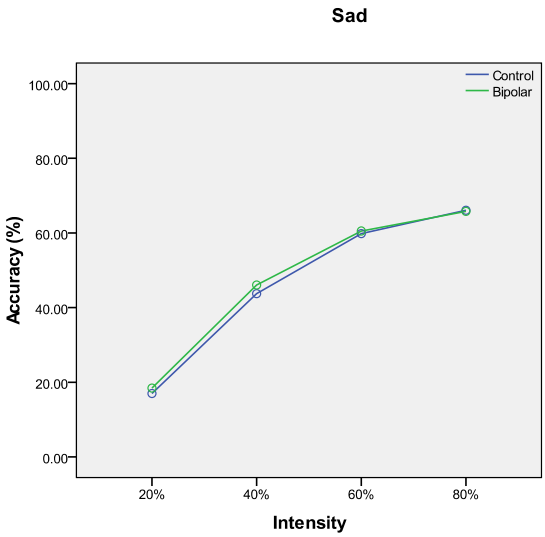


Figure 8.9e: Performance on the facial expression recognition test. Mean % accuracy at each level

Table 8.34: Results of the facial expression recognition task. Means and standard deviations of % correct at each intensity level for each emotion.

	Control (n=27)		Bipolar (n=38)	
	mean	s.d.	mean	s.d.
Angry				
Correct 20%	9.82	12.43	10.53	14.97
Correct 40%	31.25	22.18	32.24	20.06
Correct 60%	63.39	31.54	50.66	24.31
Correct 80%	78.57	24.26	70.39	23.86
Correct Total %	45.78	16.66	40.95	13.39
Disgust				
Correct 20%	6.25	14.63	5.26	11.85
Correct 40%	45.54	24.58	40.13	26.98
Correct 60%	76.79	22.49	65.79	29.88
Correct 80%	76.79	25.39	71.71	22.64
Correct Total %	51.35	14.96	45.73	15.91
Fear				
Correct 20%	10.71	15.85	16.45	20.37
Correct 40%	66.07	26.54	63.82	20.71
Correct 60%	83.93	20.65	76.97	27.50
Correct 80%	84.82	21.88	80.92	21.30
Correct Total %	61.40	15.22	59.54	14.73
Happy				
Correct 20%	35.71	26.73	46.71	28.58
Correct 40%	85.71	18.54	84.87	21.39
Correct 60%	93.75	12.95	95.39	11.41
Correct 80%	95.54	11.89	97.37	7.78
Correct Total %	77.70	12.45	81.09	13.36
Sad				
Correct 20%	16.96	18.07	18.42	18.09
Correct 40%	43.75	26.02	46.05	32.64
Correct 60%	59.82	26.65	60.53	25.09
Correct 80%	66.07	29.04	65.79	23.55
Correct Total %	46.67	18.95	47.71	16.86
Neutral				
Correct Total %	81.25	23.20	76.32	25.30
Void	0.21	1.13	0.26	0.76

VOCAL EXPRESSION RECOGNITION TASK

The results of the vocal emotion recognition task are shown in Table 8.35 below. The repeated measures ANOVA indicated no main effect of group ($F_{1,64}$=0.23, p=0.64), a significant main effect of emotion ($F_{6,384}$=33.68, p<0.001), and no significant interaction effect ($F_{6,384}$=0.42, p=0.87).

Table 8.35: Performance on the vocal expression recognition task. Mean correct responses for each emotion (max = 5)

	Control (n=28)		Bipolar (n=38)		
	mean	s.d.	mean	s.d.	d
Angry	2.43	1.07	2.26	1.11	0.15
Disgust	2.50	1.00	2.21	1.07	0.28
Fear	1.89	1.40	1.95	1.09	-0.04
Happy	2.75	1.32	2.47	1.35	0.21
Sad	3.57	1.26	3.63	0.85	-0.06
Surprised	3.75	1.27	3.84	1.10	-0.08
Neutral	3.57	1.07	3.61	0.89	-0.03

Analysis of the responses individuals made and how often they used each of the category labels revealed that the two groups used the seven emotion labels similarly (Table 8.36; χ^2_6 =6.35, p=0.39)

Table 8.36: Percentage of times each emotion label was used throughout the task by each group (including correct and incorrect responses)

	Control	Bipolar
Angry	9	9
Disgust	14	11
Fear	10	10
Happy	11	11
Sad	19	20
Surprised	16	18
Neutral	20	22

FACIAL DOT PROBE

The results of the dot probe task are reported below in Table 8.37. There were no significant differences between the groups in engagement or disengagement for any of the emotions (all $p>0.05$), although the difference between the groups for engagement with sad faces was marginally significant ($p=0.062$). For the sad faces, the bipolar patients had a higher engagement score than the controls (this was true for each of the emotions, although none of the other differences approached statistical significance) and it was the only emotion for which the engagement scores had different signs in the two groups (positive in the patient group and negative in the control group). Positive engagement scores indicate that participants showed enhanced engagement with the emotional stimulus, i.e. they were quicker to respond to valid trials containing a sad face than those containing a neutral face. This could be taken to suggest that sad faces captured attention of the bipolar patients moreso than the neutral faces. The controls showed the opposite pattern, showing a mean negative engagement score for sad faces.

The only difference on the task which reached statistical significance at conventional levels was the number of errors – patients with bipolar disorder made significantly more errors than the controls on the task ($t_{59}=-2.16$, $p=0.036$).

Table 8.37: Engagement and disengagement indices of the dot probe task.

	control (n=28)		Bipolar (n=33)				
	mean	s.d.	mean	s.d.	t_{59}	p	d
Engagement (neutral – emotion; valid trials)							
Anger	-8.23	34.54	-0.11	52.74	-0.70	0.488	-0.18
Disgust	2.24	49.87	7.42	43.01	-0.44	0.665	-0.11
Happy	-20.48	44.62	-7.06	68.11	-0.89	0.376	-0.23
Sad	-6.23	36.51	13.29	42.57	-1.90	0.062	-0.49
Disengagement (emotion – neutral; invalid trials)							
Anger	0.61	35.76	-0.53	38.42	0.12	0.905	0.03
Disgust	1.24	49.77	-3.64	42.42	0.41	0.680	0.11
Happy	11.57	60.52	1.84	47.69	0.70	0.485	0.18
Sad	12.75	58.25	2.16	43.41	0.81	0.420	0.21
Errors	0.50	0.79	1.21	1.69	364.0[a]	0.113	-0.53

[a] Mann-Whitney U

For the recognition task associated with the dot probe task, there were no significant differences between the groups in hits, misses, false alarms, or correct rejections for any of the emotions (data not shown). The difference between correct responses (hits + correct rejections) for happy faces almost reached statistical significance (patient mean = 7.42±1.73, control mean = 8.29±2.02, t_{59}=1.8, p=0.078), which on further inspection was due to fewer correct rejections (which is the reciprocal of more false alarms) by the patient group (patient mean = 4.06±1.46, control mean = 4.71±1.33, t_{59}=1.82, p=0.074).

SEMANTIC STROOP

The results of the semantic stroop task are shown in Table 8.38 below. The average reading times for the different categories of words indicate that the patient group were slower than the control group for all three categories of words, however

differences did not reach significance at conventional levels (negative and neutral words showed a trend towards significance, p=0.076 and p=0.075 respectively).

For the error scores, the bipolar patients made significantly more errors than control participants on the neutral words (t_{32}=-2.54, p=0.016), and the difference between errors on the positive words showed a trend towards statistical significance (t_{47}=-1.79, p=0.080).

There were no significant differences between the groups for the relative differences between the neutral words and either the negative words or the positive words (p>0.05).

Table 8.38: Semantic Stroop performance. Mean reading times and interference score for each category of words.

	Control (n=27)		Patient (n=33)					
	mean	s.d.	mean	s.d.	t	df	p	d
Reading time (average ms/word)								
Negative	729.50	128.36	811.69	205.25	352.0[a]	-	0.165	-0.47
Positive	731.41	131.34	793.27	183.52	-1.47	58	0.147	-0.38
Neutral	710.60	109.01	782.33	180.00	356.0[a]	-	0.184	-0.47
Errors								
Negative	0.19	0.40	0.30	0.47	-1.06	58	0.294	-0.27
Positive	0.15	0.36	0.42	0.79	-1.79	47	0.080	-0.43
Neutral	0.00	0.00	0.36	0.82	-2.54	32	0.016	-0.60
Neutral minus negative	-20.68	53.80	-29.36	53.68	0.62	57	0.540	0.16
Neutral minus positive	-21.08	48.83	-10.94	61.37	-0.69	57	0.494	-0.18

[a] Mann-Whitney U

Within the patient group, reading time for depression-related words was significantly slower than reading time for neutral words (paired t-test, t_{33}=3.14, p=0.004) and reading time for mania-related words was not significantly different from reading

time for neutral words (t_{33}=1.02, p=0.31). In the control group the opposite pattern was the case. Reading time for depression-related words was not significantly different to reading time for neutral words (t_{27}=1.83, p=0.078) and reading time for mania-related words was significantly slower than reading time for neutral words (t_{27}=2.26, p=0.033).

READING THE MIND IN THE EYES TEST

There were no significant differences between patients and controls for this task (patient mean (s.d.) = 26.69 (4.03), control mean (s.d.) = 26.79 (3.5), t_{62}=0.1, p=0.93).

DISCUSSION

The results of the emotion processing tasks in general indicated no impairment in the patients with bipolar disorder. There were no differences between the groups in facial expression recognition (assessed using either the facial expression recognition task or the Reading the Mind in the Eyes test) or in vocal emotion recognition. In the facial dot probe task, patients with bipolar disorder made significantly more errors than controls, but otherwise performed similarly showing only a trend towards greater engagement with sad faces than control participants. The semantic Stroop paradigm indicated that participants with bipolar disorder were generally slower than controls, most notably for the neutral words and the negative words (although these differences were only trends). Once again, patients made significantly more errors than controls, which reached statistical significance in the neutral condition and was a trend for the positive words. However, there were no significant differences in the Stroop effect, i.e. the level of interference, for either the mania-related or the depression-related words. Within the groups, patients showed evidence of significantly slower reading times for negative words compared to neutral words but no significant difference in reading times

for mania-related words. The opposite was true for the control group who were significantly slowed for the mania-related words, but not for the depression-related words (although this latter difference showed a trend towards significance).

The finding of no differences in explicit emotion recognition is in line with one previous study reporting no differences in facial or vocal emotion identification in euthymic patients with bipolar disorder (Vaskinn et al., 2007). Other studies have reported enhanced recognition of disgust (Harmer et al., 2002), enhanced recognition of fear (although this was in a sample of only 8 patients all of whom had bipolar II disorder and recognition was enhanced relative to other patient groups, not relative to controls) (Lembke and Ketter, 2002), and impaired recognition of fear (Venn et al., 2004). There has only been a small number of studies and the inconsistent findings imply that any deficit in emotion identification is at best subtle and does not affect all emotions or all modalities. However, it is difficult to conclude anything with confidence, given the differences in the methodologies used to date.

The vocal emotion paradigm employed was new and designed specifically for the present study. The fact that no differences were evident on this task may reflect that there are no differences there to find. However, it is also possible that without more extensive task-development work it was not a valid measure of vocal emotion recognition. The performance of both groups varied greatly across the different emotions. Specifically fear was relatively poorly recognised, whereas surprise was recognised relatively well. This may reflect the accuracy with which the stimuli captured the relevant emotion, or it may reflect the fact that agreement for vocal emotion is generally low – certainly lower than that observed for facial expressions (Scherer, 2003). However, fear has been identified as one of the emotions reliably recognised from vocal cues (Pittam and Scherer, 1993). Despite any concerns about the validity of the test, if

patients had a deficit or bias in their interpretation of the stimuli, the overall pattern of responses by each of the groups may still have differed. For example, biased responding (perhaps more likely to be revealed in circumstances when a correct answer is not obvious) would be reflected in consistent over- or underuse of a specific label. The two groups showed no significant differences in their overall use of the different category labels, indicating no evidence of a bias. It is still possible in these circumstances that patients' biases cancelled one another out, whereas the control group showed a consistently similar pattern of responses. There are two issues with this possibility. One is that if patients' biases cancelled one another out, it still implies that differences in responding are due primarily to individual differences between patients rather than something they all had in common that is 'caused by' having bipolar disorder. The second is that such a situation would be reflected in different variances between the groups, with the controls showing a lower amount of variation than patients. In fact, the variance between the groups was very similar (only showing statistically significant differences on Levene's F-test for two of the emotions, one indicating significantly higher variance in the control group and one in the patient group) making this explanation unlikely.

The results of the semantic Stroop paradigm partially replicated findings in manic patients. Earlier studies have shown evidence of increased interference from depression-related words in patients with manic symptoms (and in an analogue sample selected for high hypomanic traits) (Bentall and Thompson, 1990, Lyon et al., 1999), which the authors took as evidence that manic symptoms arise as an attempt to avoid the extreme negative feelings associated with the depressed state. There was no significant between group difference in the interference effect from depression- or mania-related words in the present study. However, within the groups the results broadly showed the same as has been found in manic patients – patients with bipolar disorder are significantly slower

to name depression-related words than neutral-related words, but the same is not true of mania-related words. This may reflect an underlying vulnerability to mood episodes in general, rather than specifically indicating risk for mania. It would be necessary to follow patients up longitudinally to identify whether this bias is associated with mood episodes of a specific polarity or whether it is a general bias observed in all phases of the illness.

The fact that there was no between-group effect on the interference scores may be a question of power and the differences would have reached significance in a larger sample. Alternatively, it may relate to the way the task was presented. Previous versions of the task have used a card-Stroop, presenting the semantic categories one at a time with a card-full of words belonging to the same category (Bentall and Thompson, 1990, French et al., 1996, Kerr et al., 2005, Lyon et al., 1999). McKenna & Sharma (2004) suggested that this confounds two potential routes by which the Stroop may operate – a fast route whereby the stimulus exerts an immediate interference effect, and a slow route whereby continued presentation of words of the same category has a cumulative effect as the semantic information is more fully processed over time (McKenna and Sharma, 2004). The present study presented words singly and recorded reaction time for each word. In the semantic condition, the three categories of words were mixed together and presented in random order, thus preventing the operation of the 'slow' route. It may well be that any Stroop effect operates via this alternative mechanism, hence the weaker evidence in the present study.

{New Para} Indeed, this is concordant with the findings of McKenna & Sharma (2004), who reported that there was very little evidence that the fast route contributed to the overall emotional Stroop effect. Additionally, findings in patients with major depression have indicated that processing biases are more likely to emerge in situations when the stimuli are self-relevant or processed in such a way as to assess their relation

to the self. For example, in incidental learning paradigms, patients with major depression show a recall bias for negative interpersonal trait adjectives, especially if during the list presentation items were encoded in relation to the self (e.g. 'typical of me' or 'not typical of me') (Mogg and Bradley, 2005). In relation to the present study, the depression- and mania-related words were not idiographically selected (i.e. the stimulus list was the same for each participant) and participants were not required to process the words in relation to the self. It may be that either of these modifications would increase the likelihood of a Stroop effect emerging.

{New para}There is also the possibility that the three categories of words were not sufficiently distinct. The words were rated by independent raters to identify whether people associated each word with depression, mania or neither. The average rating for each group of words as a whole (mania-related, depression-related, and neutral) indicated distinct differences between the groups. However, the mean ratings for each individual word showed there was some minor overlap between the categories. This would need to be improved in future studies.

{New para}Finally, patients with bipolar disorder made more errors on the task and, given that only the reaction time for correct responses was analysed, it may be that the Stroop effect noted in other studies is generated by the participant making more errors rather than responding more slowly due to increased interference. In the standard card Stroop, errors tend to be reflected in a slower total time as participants often attempt to correct their error. In the present study, self-corrections were common, however as the voice-activated response key had been triggered by the incorrect response, it was not possible to identify the time at which the correct response started. As such, it was necessary to discard errors before the response time data were analysed. Likewise, the data was cleaned before response times were analysed to exclude

unfeasibly fast responses (which, for example, were sometimes due to the subject breathing in sufficiently loudly to trigger the voice key), or exceptionally slow responses (those that were three standard deviations above the participants own individual mean reaction time for each category). Although the aim was to exclude the influence of outliers on the results, it is possible that unusually slow responses were a reflection of the Stroop effect in operation and the upper limit should have been removed. Even so, there were no significant differences between the patient and control groups in the number of excluded values for any of the semantic categories, so this effect is like to be very minor.

The facial dot probe task by and large indicated no differences in attentional engagement with emotional faces in patients with bipolar disorder. Earlier evidence of a reduced bias to positive stimuli in euthymic bipolar patients had been demonstrated using word stimuli in a standard dot probe task (i.e. with a dual-cue (Jongen et al., 2007)). Previous studies in symptomatic patients using the single-cue paradigm with facial stimuli had demonstrated increased engagement with angry faces in depressed patients (Leyman et al., 2009, Leyman et al., 2007). The present findings did not indicate any differences in engagement with angry faces, although there was a trend towards greater engagement with sad faces in the patient group. Given the small sample size, this finding needs to be replicated in a larger sample before any conclusions can be drawn. The stimulus presentation time used in the current study (500ms) may not be the optimal time for identifying attentional bias effects. Patients with anxiety disorders show cognitive biases with very short presentation times, whereas depressed patients are more likely to show biases when stimulus presentation exceeds 1000ms (Mogg and Bradley, 2005). Exploring both shorter and longer presentation times (as has been done in studies of symptomatic patients (Leyman et al., 2009) may be more informative about the nature of any attentional bias to emotional stimuli in patients with bipolar disorder.

Additionally, in the present task participants were instructed that there would be a recognition task at the end of the attentional task and they would be asked if they could identify any of the faces they saw during the task. Participants were therefore encouraged to look at the faces. This instruction was designed to prevent participants 'zoning out' or simply ignoring the facial stimuli in order to respond as accurately as possible to the following probe. However, it may have created interference effects or increased the burden on the participants who may have tried to consciously remember the faces. Patients made more errors on the dot probe task than controls. This may be due to general attentional impairments, which are found in patients with bipolar disorder (Clark et al., 2002). Even though attempts were made to minimise the length of the task (the task length was kept below 5 minutes) and a break was included to minimise performance impairment resulting from poor sustained attention, patients still made more errors. Any effects on attention due to the emotional stimuli may have been outweighed by general impairment in attention and it may be that this task paradigm is not suited to exploring emotional biases in this population. One final point relating to the dot probe task is that the stimuli were presented in random order. As discussed for the Stroop task above, effects of emotional stimuli may build over repeated presentations (the 'slow' route as proposed by Mckenna & Sharma 2004). Presenting the stimuli in blocks of a given emotion may be a valuable extension to ascertain whether repeated presentation of the same emotional expression has cumulative effects on performance. Whether the effects demonstrated would still be operating via attentional mechanisms would then need to be explored.

The present findings have added to the small body of literature examining emotion-processing in euthymic patients with bipolar disorder. The overall conclusion is that there is very little evidence of abnormalities in the aspects of emotion-processing examined here. Future studies should build on these findings, improving on some of the

methodological issues discussed above, in order to build a more complete picture of emotion processing in euthymia and how it differs to emotion processing during episodes. On the face of it, the present results do not support the hypothesis that biased emotion processing remains as a possible vulnerability factor for repeated episodes in patients with bipolar disorder.

CHAPTER 9: PSYCHOSOCIAL FUNCTIONING IN BIPOLAR DISORDER

INTRODUCTION

In Chapter 4 the problem of functional impairment in euthymic patients with bipolar disorder was highlighted. A large proportion of patients in a non-acute phase of illness continue to experience restrictions in the degree to which they participate in everyday life (MacQueen et al., 2001). After an episode many fail to reach the level of functioning seen in the general population, or fail to return to the level of functioning experienced prior to the onset of their illness (Tohen et al., 2000).

The previous chapters have highlighted ways in which euthymic patients with bipolar disorder may process information differently to those without the disorder. Particularly, deficits were noted in aspects of executive functioning and verbal memory. The aim of the present chapter is to investigate whether these impairments have an association with psychosocial function.

Research to date investigating predictors of social functioning in bipolar disorder has been limited by a number of factors. The measurement instruments used have either been very general, or have confounded symptoms with function, or have not been designed for or validated in people with mood disorders. There has been a greater focus on clinical and sociodemographic factors, whereas cognitive function and social information processing have been relatively neglected. In studies using regression analyses, the variance explained has tended to be fairly low (i.e. around 30%) suggesting that key explanatory factors may not have been investigated.

The aim of the present chapter is to investigate social functioning in euthymic patients with bipolar disorder using a broad selection of measures of functioning – including both clinician-rated and participant-rated measures, as well as measures derived specifically for populations with mood disorders. The second more exploratory aspect of the analyses aims to investigate the relationship between social function, cognitive function, emotion processing and other factors using regression analysis.

METHODS

PARTICIPANTS

As described previously in Chapter 5, page 97.

MATERIALS

Copies of the psychosocial function scales are reproduced in Appendix 7 on page 324.

CLINICIAN-RATED MEASURES:

1. GLOBAL ASSESSMENT OF FUNCTIONING (GAF) (APA, 1994)

The GAF is included as Axis 5 in the Diagnostic and Statistical Manual of Mental Disorders 4th edition (DSM-IV) (APA, 1994). It is a 100-point scale ranging from 1 to 100 designed to measure level of functioning. The scale is grouped into 10 broad categories with scores between 1-10 representing the lowest level of functioning and scores between 91-100 representing the highest level of functioning. Although banded in this way, the bands only serve as a rough guide and within each band a higher score represents better functioning. The score can take any whole number between 1 and 100. The scale measures functioning in its broadest sense and attempts to capture social, relational and occupational aspects of functioning as well as level of symptoms. In the case where an individual's functioning is not the same for each of these aspects, they are

rated in accordance with the lowest. The scale has been reported to have 'modest' reliability and validity (Goldman et al., 1992a).

2. Specific Levels of Functioning (SLOF) (Schneider and Struening, 1983)

The SLOF is a 43-item assessment of behavioural functioning and daily living skills. The items are divided into six domains: physical functioning (5 items, e.g. use of arms, legs, sensory difficulties), personal care skills (7 items, e.g. grooming, ability to feed self, dress appropriately, general self-care), interpersonal relationships (7 items, e.g. forming friendships, communication, accepting contact with others, behaviour in groups), social acceptability (7 items, e.g. conforming with social norms, violence towards self or others), activities (11 items, e.g. ability to travel independently, manage finances, run a household, use of leisure time), and work skills (6 items, e.g. employable skills, punctuality, ability to work independently with only appropriate supervision). Each item is scored from 1 (lowest level of functioning) to 5 (highest level of functioning). The score for each subscale is derived by summing the scores for the individual items, and a total score is derived by summing the scores for all items (maximum score 43x5=215). The scale has been reported to show good inter-rater reliability when used on patients in community settings (intra-class correlation coefficient (ICC)=0.62), good factoral validity (the same factor structure of the items was demonstrated in four different samples), and four of the six factors showed excellent internal consistency (Chronbach's $\alpha \geq 0.92$) with the remaining two showing satisfactory internal consistency (physical functioning $\alpha=0.76$, social acceptability $\alpha=0.80$).

3. Longitudinal Follow-up Evaluation Range of Impaired Functioning Tool (LIFE-RIFT) (Leon et al., 1999)

The LIFE-RIFT is a brief rating scale designed to assess functioning specifically in patients with affective disorders. It assesses four domains: employment, interpersonal

relations, satisfaction and recreation. The employment domain is subdivided into 3 areas – work, household and student. The individual is rated on each of the applicable areas (those not applicable are rated zero) on a six point scale from 1 to 6. A score of one represents no impairment and five represents severe impairment (six represents insufficient information). The overall domain score for employment is the highest-rated sub-section of the three (i.e. the lowest level of functioning). The interpersonal relations domain is subdivided into 4 areas – relations with spouse, children, other relatives, and friends. Each is rated on the same six-point scale as that used for the employment domain. The overall domain score is the highest rating of each of these four subscales. Both the satisfaction and recreation subscales are each rated on a single six-point scale with the same properties as those used to rate the other domains. The overall score for the entire scale is simply the sum of the four individual domain scores (range 4-20). The scale shows good internal consistency reliability (Chronbach's α 0.81-0.83) and excellent inter-rater reliability (ICC = 0.94). The scale has been validated specifically in a sample of patients with bipolar disorder and again showed good internal consistency reliability (Chronbach's α between 0.78-0.84) and excellent inter-rater reliability (ICC = 0.94)(Leon et al., 2000).

PATIENT-RATED MEASURES:

4. MEDICAL OUTCOMES SURVEY SHORT FORM (36-ITEM) (SF-36)(WARE AND SHERBOURNE, 1992) RAND SCORING (HAYS ET AL., 1993)

The SF-36 is widely used as a measure of health-related functioning. It has 36-items which form 8 subscales: physical health (10 items), role limitations due to physical health (4 items), social functioning (2 items), pain (2 items), emotional well-being (5 items), role limitations due to mental health (3 items), energy (4 items), and general perception of health (5 items). The final item is a change score, reflecting a comparison

between current state of health and state of health one year ago. As this was not relevant for the present study, this item was not scored. The items are rated on different scales with some having binary response options (role limitations due to physical health and role limitations due to mental health), some three response options (physical health scale), some 5 response options (social functioning, pain and general perception of health) and others 6 response options (energy and emotional well-being). The scale options usually reflect frequency or degree of impairment. Several items are reverse-scored.

Owing to the differences in item numbers per subscale and rating options per item, the RAND scoring scale was used (Hays et al., 1993). This rescales the scores for each item, converting them to scores between 0 (worst possible function) and 100 (best possible function). For the binary response items, responses are therefore scored either 0 or 100, for the three-response items they are scored either 0, 50 or 100, for the five-response items they are scored either 0, 25, 50, 75 or 100, and so on. The subscale scores are derived by summing the scores for each item belonging to that subscale and dividing by the number of subscale items answered (range = 0 to 100). A total scale score is derived by summing the rescaled scores for all items answered and dividing by the total number of items answered (range = 0 to 100). Higher scores indicate better functioning. The psychometric properties of the SF-36 have been explored extensively and it has been demonstrated to show good divergent validity for physical and mental health-related functioning (McHorney et al., 1993) and reliability coefficients ranging from moderate to good across the different subscales (between 0.65-0.94) (McHorney et al., 1994).

5. LIFE FUNCTIONING QUESTIONNAIRE (LFQ) (ALTSHULER ET AL., 2002)

The LFQ was designed as a self-report measure specifically for patients with affective disorders. The essence of the scale contains 14 items and assesses function in four domains: leisure activities with friends, leisure activities with family, duties at home, and duties at work (or school, or activity centre, as applicable). The two leisure-related domains are each rated on three dimensions – the amount of time spent with friends/family, the degree of conflict, and the amount of enjoyment. Each of these three dimensions is rated on a four-point scale where 1 represents no problems and 4 represents severe problems. The duty-related domains are each rated on four dimensions – the same three mentioned above plus the quality of work. Both are rated on the same four-point scale. Domain scores are derived by summing the scores on each of the dimensions and dividing by the number of dimensions rated. No overall score is calculated. The scale has been shown to have good test-retest reliability (ICC ≥0.7 across subscales), and good internal consistency (α≥0.84 across subscales), as well as good concurrent validity showing modest to strong correlations with another self-rated functional outcome measure (r ranging between 0.57 and 0.86) (Altshuler et al., 2002).

6. SOCIAL ADJUSTMENT SCALE – SELF REPORT VERSION (SAS-SR) (WEISSMAN, 1999, WEISSMAN ET AL., 1978)

The SAS-SR is a 54-item scale which is divided into 8 subscales: work (for pay; 6 items), housework (unpaid; 6 items), education (completed only by students; 6 items), social and leisure (11 items), family outside the home (8 items), primary relationship (9 items), parental (4 items), and family unit (4 items). Each item is rated on a 5-point scale where 1 indicates good functioning and 5 indicates poor functioning. The first three scales (work, housework, and education) are rated only to the extent that they apply, and only one of the three (whichever is identified as the rater's primary role) is scored.

The scores for each subscale are calculated by summing the score for each item and dividing by the number of items answered. The total score is calculated in the same way. The scale has shown high internal consistency ($\alpha>0.7$) and good test-retest reliability ($r>0.7$) (Weissman, 1999).

PROCEDURE

The clinician-rated measures were conducted by a psychiatrist. For the patient group, the assessment of social functioning was conducted on the day that they completed the neuropsychological tests and took place at the same time as the clinician-rated mood measures (that were used to establish continued euthymia). For the controls, social functioning was assessed at the screening session as no further clinical assessments were conducted with the control participants prior to their neuropsychological tests. For most of the control participants the gap between screening and testing was a week or shorter.

DATA ANALYSIS – BETWEEN-GROUP COMPARISONS

Data were analysed using SPSS version 17.0. Scores on social functioning measures were first plotted using a box plot to identify variables with notable outliers (values more than three times the height of the box (which represents the inter-quartile range)). The two groups (patients and controls) were compared on the functioning measures using independent samples t-tests with $p\leq0.05$ set as the significance level. For variables showing extreme outliers on the boxplots, an equivalent non-parametric independent samples test was conducted (the Mann-Whitney U test), and where the results differed in terms of statistical significance, the Mann-Whitney test result was reported. As discussed previously, formal tests of normality were not used before using parametric techniques as the sample size was deemed to be too small for these tests to

be reliable and parametric tests are relatively robust to deviations from normality in the absence of outliers.

Effect sizes for these measures were calculated using Cohen's d where d=(control mean-patient mean)/pooled standard deviation. Effect sizes for indices where a higher score indicated worse functioning were converted (multiplied by -1) such that a positive effect size always indicates poorer functioning in the patient group. Conventions for Cohen's d are that $0.2<d<0.5$ is a small effect, $0.5\leq d<0.8$ is a medium effect, and $d\geq 0.8$ is a large effect.

In order to explore the relationship between the different functioning measures, the correlations between them were computed. The functioning domains that the clinician-rated questionnaires had in common (e.g. work functioning, interpersonal functioning, etc.) were correlated with one another using Pearson correlations. The same procedure was repeated for the functioning domains that the self-rated measures had in common.

DATA ANALYSIS – PREDICTING FUNCTION

Forwards stepwise regression analyses were used to explore the relationship between functioning and cognitive function, emotion processing, and psychological factors (such as self-esteem, mood, dysfunctional attitudes). The analyses were exploratory in nature and proceeded through several stages in order to identify the model(s) with the best explanatory power.

DATA ANALYSIS – IDENTIFYING DEPENDENT VARIABLES

In the first step the dependent variables were selected. The aims were to investigate whether :- 1) self-rated and clinician-rated functioning measures are predicted by the same indices as one another. One measure of clinician-rated and one

measure of self-rated function was selected. The LIFE-RIFT was selected as the clinican-rated measure as it was specifically designed for use in mood disorder and showed stronger evidence of convergent validity with the GAF (unlike the SLOF, which correlated poorly with both the GAF and the LIFE-RIFT; see Table 9.40 on page 242). The SAS-SR was selected as the self-rated measure, as it was the most comprehensive measure used. All the self-rated measures showed a similar degree of correlation with one another (see Table 9.42 on page 245) therefore no single measure stood out at more appropriate than the others on convergent validity grounds. Instead, the SAS-SR contained more subscales in common with the clinican-rated measure and therefore was more readily comparable. An additional clinician-rated measure, the GAF, was selected as a global functioning measure for comparability with previous studies. 2) whether different aspects of functioning are predicted by different indices. Three specific areas or domains of functioning were selected as dependent variables – work, interpersonal relationships, and recreational functioning. These three areas were selected for a number of reasons. They are *relatively* concrete and readily-understandable areas, do not include quality of life-related aspects such as life satisfaction, and were indices which the two selected functioning measures (LIFE-RIFT and SAS-SR) had in common (although note that only the LIFE-RIFT has a single interpersonal subscale; the SAS-SR divides interpersonal relationships across several categories of relationship (family, primary relationship, and parental). As different numbers of participants answered these sections depending on which were applicable, it was not possible to amalgamate these scores into an overall interpersonal functioning subscale score). Additionally these three functioning areas were relevant for the majority of the sample therefore these questions were answered by most of the participants which maximised the available sample size. The total scale scores on the LIFE-RIFT and the SAS-SR were also included as outcome variables, as was score on the GAF (which provides only a total score and no subscale scores). This process

provided eight dependent variables – a clinician-rated and self-rated measure of work functioning (LIFE-RIFT work subscale and SAS-SR work subscale respectively), a clinician-rated measure of interpersonal functioning (LIFE-RIFT interpersonal subscale, a clinician-rated and self-rated measure of recreational functioning (LIFE-RIFT Recreation subscale and SAS-SR Social & Leisure subscale respectively), and two clinician-rated and one self-rated measures of total functioning (GAF, LIFE-RIFT Total and SAS-SR Total respectively). A separate regression equation was estimated for each dependent variable.

DATA ANALYSIS – IDENTIFYING CANDIDATE PREDICTORS: COGNITIVE MEASURES

The second step of constructing the models involved identifying candidate predictors. Neurocognitive test scores were selected as potential predictors on the following grounds. They showed a statistically significant difference between patients and controls ($p<0.05$) and when more than one index of a given test showed a significant difference, only the index showing the largest effect size was selected. These criteria were chosen for two reasons. The first was to use predictors that showed differences between patients and controls. Although it is feasible (and indeed likely) that a variable showing no difference between the groups nonetheless relates significantly to functioning, the relationships of most interest to understanding the reasons behind poor functioning are those with variables on which patients with bipolar disorder show impairment. The second reason was to minimise the chances of strong correlations between the independent variables (which may be more likely when drawing more than one outcome variable from the same test), as this could cause unstable estimation in regression due to multicollinearity. This process resulted in 12 potential predictor variables: verbal fluency rule breaks ($VerFlu_{RULEBREAKS}$), category fluency total ($CatFlu_{TOTAL}$), Hayling Sentence Completion Test category B errors ($HSCT_B$), Rey Auditory Verbal Learning Test trial A1 ($RAVLT_{A1}$), Wechsler Memory Scale Logical Passages Test total story

units ($WMS_{LP\text{-}TOTAL}$), Behavioural Assessment of the Dysexecutive Syndrome Zoo Map subtest version 1 raw score ($BADS_{ZM\text{-}V1}$), total score for the second administration of the Wechsler Adult Intelligence Scale Digit Symbol Substitution Test ($DSST_2$), total time for Trail Making Test part A (TMT_A), number correct of the easy trials at the 0-second delay on the Simultaneous and Delayed Match to Sample test ($SDMTS_{0EASY}$), total errors on the Self-Ordered Pointing Test ($SOPT_{TOTAL}$), errors on the Color-Word trial of the Stroop test ($Stroop_{ERRCW}$) and total trials administered on the Wisconsin Card Sorting Test ($WCST_{TOTALTRIALS}$).

DATA ANALYSIS – IDENTIFYING CANDIDATE PREDICTORS: EMOTION MEASURES

Potential predictors of functioning drawn from the emotion-processing variables were selected in a similar manner. However, as there were very few statistically significant differences between patients and controls on any of the emotion processing measures, only measures showing an effect size $d \geq |0.4|$ were used. This resulted in 9 potential predictors: three measures from the facial expression recognition test – correct responses for angry faces displayed at the 60% level of intensity ($Angry_{60}$), happy faces displayed at the 20% level of intensity ($Happy_{20}$), and disgusted faces displayed at the 60% level of intensity ($Disgust_{60}$), three measures from the dot probe test – engagement with sad faces ($DotProbe_{ENGAGESAD}$), dot probe total errors ($DotProbe_{ERROR}$), and recognition of happy faces ($DotProbe_{RECOGHAPPY}$), and three measures from the semantic Stroop – reading time for negative words ($SemStroop_{RTNEG}$), reading time for neutral words ($SemStroop_{RTNEU}$), and errors for positive word ($SemStroop_{ERRPOS}$). No indices from the vocal emotion recognition test reached the criterion for entry. Two further predictors were added to capture general emotion recognition processes, primarily for face validity purposes – total correct on the facial expression recognition test ($Facial_{TOTAL}$) and total correct on the vocal emotion recognition test ($Vocal_{TOTAL}$).

DATA ANALYSIS – IDENTIFYING CANDIDATE PREDICTORS: PSYCHOLOGICAL MEASURES

The next step involved identifying psychological predictors of functioning. In a similar manner, questionnaire measures that showed differences between patients and controls were identified. Five variables were made available to the model: self-rated depressive symptoms on test day (Depression), self-rated manic symptoms on test day (Mania), score on the Rosenberg self-esteem scale (SelfEsteem), total score on the Dysfunctional Attitudes Scale (DAS_{TOT}), and score on the Trait Anxiety Inventory (TraitAnx). Ratings on the mania rating scale were made available to the model despite showing no significant differences between groups for face validity reasons. Self-rated mood symptoms were used rather than clinician-rated measures for two reasons – self-rated measures were available for both groups on the test day (whereas clinician-rated measures were only available on the screening day for the control participants), and self-rated measures had no constraints on range. The clinician-rated measures all fell below the designated cut-offs used for euthymia, however no such restriction applied to the self-rated scales.

DATA ANALYSIS – REGRESSION ANALYSIS PART I

Two sets of regression equations were estimated, one each for the patient and control group separately. For each of the eight dependent variables an equation was estimated using ordinary least squares with forwards stepwise entry. Each of the variables identified via the process described above (n=27) was made available to the model. This method uses the correlations between the dependent and independent variables to identify the strongest predictor variable first, which is then entered into the model. If this is a significant predictor (at the $p<0.05$ level), any variance not explained by that first predictor is then 'made available to' the next strongest predictor, and it is added into the model. If the added variable is a significant predictor (at the $p<0.05$ level)

then a further variable is added and so on until a predictor variable is entered that does not significantly associate with the dependent variable and does not increase the explanatory power of the model significantly. If, after the entry of variable, a previously-entered variable is reduced to non-significance ($p>0.1$), this variable is removed.

The forward entry method was chosen for two reasons. The first was in accordance with the exploratory nature of the analysis – there were no hypothesized relationships and thus a data-driven approach was selected. Once the candidate independent variables had been selected, there were no constraints as to what order the variables should enter the model. The second was that an alternative data-driven approach, the backwards entry procedure (where all independent variables are entered simultaneously and the weakest predictors removed until no further variables can be removed without significantly impairing the explanatory power of the model), would have been inappropriate for the sample size, given the number of candidate independent variables used.

Equations were checked for evidence of multicollinearity. Any variables with a variance inflation factor (VIF) greater than 10 were considered highly multi-colinear. For each model the variance proportions matrix was examined to identify whether factors associated with small eigenvalues loaded across more than one independent variable (Field, 2000).

The process outlined above resulted in eight equations describing one aspect or measure of functioning for the control group, and eight similar equations for the patient group.

DATA ANALYSIS – REGRESSION PART II: ADJUSTING FOR DEMOGRAPHICS & ILLNESS HISTORY

The results of the above analysis were not adjusted for demographic factors in either group, or for illness history variables in the patient group. The same series of models were run again adding age, years of education and estimated pre-morbid IQ (score on the National Adult Reading Test (Nelson, 1982)) to the variables available to the models for both groups. Additionally, illness history variables (age at onset, number of episodes, and number of hospitalizations) were added to the variables available to the patient group models.

DATA ANALYSIS – FACTOR ANALYSIS

The final stage in the process involved using factor analysis to reduce the number of predictor variables to a smaller number of factors, each capturing a common underlying process in the data. The factor scores on each of these underlying components were then used in a final series of regression analyses to identify which were the best predictors of functioning. The main aim was to reduce any potential problems in the regression analyses caused by interrelationships between the independent variables. The secondary aim was to distil the large array of predictor variables into a smaller number whilst retaining as much common variance as possible. Additionally, by examining which variables load together, it may be possible to identify which broad function or process shows the strongest association with functioning.

To derive the factors, initially the three groups of variables that had entered into the regression analyses were used and each group of variables was factor analysed separately. For the factor analysis, data from both participant groups was combined.

As a first step, the correlation matrices were examined and any variables that did not correlate significantly with any other variables or with only a small number of other variables (i.e. only one or two) were excluded. The analysis was run using the principal components extraction methodology in SPSS with varimax rotation. The use of an

orthogonal rotation rather than an oblique rotation was to minimise the extent to which factors would correlate with each other. To produce factor scores that do not correlate with each other, factor scores were calculated using the Anderson-Rubin method. Both of these options were selected to minimise the likelihood of multi-collinearity in the subsequent regression analyses. The solutions were checked for sampling adequacy (i.e. the variables as a whole and each individual variable had a Kaiser-Meyer-Olkin statistic >0.5) and sphericity (using Bartlett's sphericity test). In all cases the initial rotated solution was examined and any variables that loaded poorly onto a single factor or loaded evenly onto more than one factor were excluded and the analysis re-run. For present purposes a factor loading of ≥0.5 was used as a cut-off for a moderate-to-high factor loading and values below this were excluded. This is a relatively stringent cut-off and was used because the sample size was small and to ensure that included variables accounted for at least 25% of the variance in the factor in order to produce 'strong' factors. Factors were extracted provided they had an eigenvalue >1.

The factor analyses using the variables selected for the regression analyses produced solutions that had a number of problems. For the cognitive variables, the initial solution was difficult to interpret and 6 out of the 12 variables justified exclusion on the basis of the above criteria (one for KMO<0.5, three for even loading on more than one factor, and two for no factor loadings >0.5). For the emotion-processing variables, again the initial solution was difficult to interpret, and 6 out of the 11 variables had a KMO statistic <0.5 (indicating poor sampling adequacy) and a further one variable justified exclusion on the basis of even loading across more than one factor. In both cases attempts to improve the solution revealed that the results were unstable. As a result, alternative variables were selected for the factor analyses. Measures were chosen that sampled across the whole test battery and reflected key outcome variables of the tests. This had the added benefit of allowing factors to emerge from a wider selection of

variables than simply those showing statistically significant differences between the groups (see Table 9.45 and Table 9.46 for full details of the measures selected).

DATA ANALYSIS – REGRESSION PART III

The factor scores on all of the factors that were extracted were subsequently entered used in regression analyses. The same procedure was used as above, ordinary least squares with forwards stepwise entry (with criteria of $p<0.05$ for entry and $p>0.1$ for removal), and the factor scores on all of the derived factors (n=9) were made available as potential predictors regardless of the proportion of variance they explained. As the major aim of the factor analysis was to reduce problems due to collinearity and to reduce the number of variables (rather than to identify or hone the factor structure of the three individual domains), factors that accounted for a small proportion of variance (i.e. <15%) were included provided the factor loadings were high and the communalities of their associated variables were >0.5. A separate set of models were run for the control and patient groups. In addition to the factor scores, demographic factors (age, years of education and estimated premorbid IQ) were available as potential predictors.

DATA ANALYSIS – LEARNING POTENTIAL AND FUNCTIONING

The relationship between learning potential and functioning was explored separately using Spearman correlations with $p<0.05$ indicating statistical significance. The percentage change in performance between each of the four administrations of the Digit Symbol Substitution Test (DSST) was calculated and correlated with each of the eight functioning measures used as dependent variables in the regression analyses. Additionally the linear slope of all four DSST scores was calculated for each participant and correlated with the eight functioning measures.

RESULTS

CLINICIAN-RATED MEASURES

Table 9.39: Scores on the clinician-rated social functioning measures

	Control (n=24)		Bipolar (n=36)					
	mean	**s.d.**	**mean**	**s.d.**	**t**	**df**	**p**	**d**
Global Assessment of Functioning	92.04[a]	4.41	75.00[c]	10.33	9.03	54	<0.001	2.01
LIFE-RIFT								
Work	1.16[b]	0.47	3.06	1.15	-8.90	50	<0.001	2.03
Interpersonal Relations	1.69	0.91	2.32[d]	1.03	-2.47	59	0.016	0.65
Satisfaction	1.50	0.59	2.19[d]	0.70	-3.99	59	<0.001	1.04
Recreation	1.50	0.93	2.14[d]	0.71	-3.01	59	0.004	0.79
Total	5.85	1.86	9.59[d]	2.73	-5.87	59	<0.001	1.54
SLOF								
Physical Functioning[e]	24.67	0.64	24.19	1.41	1.76	52	0.085	0.40
Personal Care	35.00	0.00	34.94	0.23	1.43	35	0.160	0.31
Interpersonal Relationships	33.46	3.26	30.53	3.94	3.14	55	0.003	0.80
Social Acceptibility[f]	34.75	0.53	34.14	1.73	1.99	44	0.053	0.44
Activities	55.00	0.00	52.06	3.61	4.90	35	<0.001	1.05
Work	29.54	1.41	26.75	3.53	4.26	49	<0.001	0.97
Total	212.42	4.74	202.61	10.82	4.79	52	<0.001	1.10

[a] n=26 [b] n=25 [c] n=38 [d] n=37
[e] This variable had significant outliers, Mann-Whitney U=377, p=0.305
[f] This variable had significant outliers, Mann-Whitney U=355.5, p=0.152

The group differences in clinician-rated functioning are reported in Table 9.39 above. The patients with bipolar disorder showed significant impairment in all aspects of functioning as measured by the GAF and the LIFE-RIFT, and four out of seven indices of the SLOF. No statistically significant impairment was noted on the physical functioning, personal care, or social acceptability subscales of the SLOF and the effect sizes were

small on each of these indices (d<0.5). Effect sizes on almost all of the indices showing significant impairment were large (d>0.8) and indicated that on average patients were scoring between one and two standard deviations above the controls.

The correlations between the common-indices of the clinician-rated measures indicated that the GAF and the LIFE-RIFT correlated relatively highly with one another, whereas the SLOF correlated poorly with either of the other two scales (see **Appendix 7**Table 9.40 below).

Table 9.40: Pearson correlation matrix of the relationship between the subscale scores of the common subscales of the clinician-rated psychosocial functioning measures

		GAF	LIFE-RIFT total	LIFE-RIFT work	LIFE-RIFT interpersonal	LIFE-RIFT recreation
LIFE-RIFT total	r	-0.75				
	p	<0.001				
	n	61				
LIFE-RIFT work	r	-0.73				
	p	<0.001				
	n	60				
LIFE-RIFT interpersonal	r	-0.46				
	p	<0.001				
	n	61				
LIFE-RIFT recreation	r	-0.45				
	p	<0.001				
	n	61				
SLOF total score	r	0.32	-0.41	-0.41	-0.14	-0.33
	p	0.013	0.001	0.001	0.289	0.011
	n	59	59	59	59	59
SLOF work skills	r	0.35	-0.42	-0.45	-0.07	-0.36
	p	0.007	0.001	<0.001	0.580	0.006
	n	59	59	59	59	59
SLOF interpersonal relationships	r	0.18	-0.29	-0.14	-0.21	-0.30
	p	0.164	0.024	0.282	0.117	0.020
	n	59	59	59	59	59
SLOF activities	r	0.41	-0.41	-0.49	-0.11	-0.23
	p	0.001	0.001	<0.001	0.413	0.085
	n	59	59	59	59	59

PARTICIPANT-RATED MEASURES

Scores on the participant-rated functioning measures are reported below in Table 9.41. Patients reported significant impairment in functioning in all indices of the SF-36. In the LFQ patients did not report significant difficulties in functioning with regard to leisure time with either friends or family, but did show impairment in both duties at home and work. On the SAS-SR significant impairment was noted in all domains, except functioning as a parent and as a family unit. In general effect sizes of the impaired domains were large ($1<d<1.5$) indicating sizeable impairment.

Overall, the results of the between-group comparisons indicate what has been found previously in this patient group – significant and relatively sizeable impairment in almost all domains of life functioning. Only physical functioning (i.e. functioning of limbs and sense organs), personal care (washing, dressing appropriately, feeding oneself etc.) and social acceptability of behaviour showed no impairment in the clinician-rated measures. This is unsurprising in a euthymic sample. In the self-rated measures only leisure activities with friends and family, and functioning as a parent and family unit showed no impairment.

Table 9.41: Scores on the participant-rated social functioning measures (n included in table as some subscales are not applicable to everybody)

	Control			Bipolar			t	df	p	d
	n	mean	s.d.	n	mean	s.d.				
SF-36										
Physical Functioning	28	93.93	6.43	38	80.74	21.20	3.61	46	0.001	0.79
Role limitations due to physical health	28	88.39	28.45	38	69.08	38.74	2.34	64	0.023	0.56
Role limitations due to emotional problems	28	97.62	8.74	38	54.38	43.44	5.97	41	<0.001	1.29
Energy	28	73.21	18.62	36	48.94	20.70	4.86	62	<0.001	1.22
Emotional well-being	28	86.43	9.57	36	66.78	17.13	5.81	57	<0.001	1.37
Social functioning	28	97.32	7.10	36	68.40	24.91	6.63	42	<0.001	1.50
Pain	28	93.30	10.65	36	80.21	22.43	3.08	53	0.003	0.72
General health	27	81.11	15.02	36	59.17	23.25	4.54	60	<0.001	1.09
Overall score	28	88.79	7.09	36	67.71	18.04	6.40	48	<0.001	1.47
LFQ										
Leisure with friends	27	1.12	0.26	33	1.32	0.47	355[a]	52	0.053	0.51
Leisure with famil	27	1.23	0.50	35	1.35	0.72	-0.71	60	0.481	0.18
Duties at home	28	1.18	0.31	37	1.75	0.67	-4.57	54	<0.001	1.04
Duties at work	25	1.18	0.33	16	1.72	0.66	-3.02	20	0.007	1.11
SAS-SR										
Work Role	26	1.24	0.25	37	1.92	0.64	-5.79	50	<0.001	1.30
Social & Leisure	28	1.75	0.61	38	2.41	0.64	-4.26	64	<0.001	1.06
Extended Family	28	1.37	0.43	36	1.73	0.72	-2.30	62	0.025	0.58
Primary Relationship	19	1.40	0.35	26	2.05	0.70	-4.10	39	<0.001	1.12
Parental	9	1.47	0.29	10	1.76	0.87	-0.93	17	0.367	0.43
Family unit	25	1.50	0.59	38	1.73	0.80	-1.22	61	0.227	0.31
Overall	28	1.48	0.31	38	2.00	0.38	-5.86	64	<0.001	1.46

[a] Mann-Whitney U

The correlations between the common-indices of the self-rated measures indicated that all three measures correlated to a similar degree with one another and the correlations were generally modest to strong (0.4-0.7; see Table 9.42 below). This indicates there is notable overlap in the information gathered by these different instruments and possibly some redundancy in measurement.

Table 9.42: Pearson correlation matrix of the relationship between the subscale scores of the common subscales of the self-rated psychosocial functioning measures

		SF36 Role limitations due to emotional problems	SF36 Social functioning	SF36 Overall score	LFQ leisure with friends	LFQ duties at home	LFQ duties at work, school, activity centre
LFQ leisure with friends	r	-0.42	-0.57	-0.39			
	p	0.001	<0.001	0.002			
	n	60	58	58			
LFQ duties at home	r	-0.59	-0.63	-0.73			
	p	<0.001	<0.001	<0.001			
	n	65	63	63			
LFQ duties at work	r	-0.66	-0.66	-0.68			
	p	<0.001	<0.001	<0.001			
	n	41	41	41			
SAS-SR work role	r	-0.57	-0.62	-0.64	0.48	0.60	0.61
	p	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	n	63	61	61	57	63	41
SAS-SR social & leisure	r	-0.33	-0.55	-0.55	0.56	0.55	0.63
	p	0.006	<0.001	<0.001	<0.001	<0.001	<0.001
	n	66	64	64	60	65	41
SAS-SR overall	r	-0.51	-0.70	-0.65	0.62	0.69	0.66
	p	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	n	66	64	64	60	65	41

PREDICTING FUNCTIONING

The results of the first stage of regression analyses are reported separately for patients (Table 9.43, page 250) and controls (Table 9.44, page 254).

PATIENTS - GENERAL

The regression equations for the patient group indicated good explanatory power, with adjusted R^2 values ranging from 33.0%-67.6%. Broadly, all of the models contained a variety of predictors, including some variables from each of the three types – psychological, cognitive and emotion-processing indices. In each of the models bar one

(self-rated recreational function), a psychological variable was the strongest predictor (explained the most variance), usually followed by a verbal memory measure (passage recall) or an executive function measure (category fluency).

In general the signs on the coefficients indicated relationships in the expected direction (i.e. poorer cognitive function/emotion-processing/psychological function was associated with worse psychosocial function). However, the signs on the coefficients for the simultaneous and delayed match to sample test index (number correct at 0-second delay), recognition of happiness at 20%, recall on trial 1 of the verbal learning test, and category B errors in the Hayling Sentence Completion Test were all in the direction indicating better performance on the test index was associated with poorer functioning. Signs in an unexpected or counter-intuitive direction can be an indicator of multicollinearity (Studenmund, 2001). However in the models affected neither the variance inflation factors nor the variance proportion matrices indicated problems stemming from multicollinearity (these issues are discussed in more depth in the discussion below).

PATIENTS - CLINICAN-RATED VS SELF-RATED MEASURES

The specific variables that entered into the models of clinician and self-rated functioning tended to vary, but the overall pattern remained similar, i.e. a psychological variable explained the largest proportion of variance, followed by a verbal memory or executive function measure, with emotion-processing variables explaining a relatively small proportion of variance. Clinician- and self-rated function were explained almost equally well in terms of adjusted R^2, apart from recreational functioning where a smaller proportion of self-rated functioning (33%) was explained than clinician-rated function (49%).

In the equations of clinican-rated function, dysfunctional attitudes featured frequently (in 3 out of 5 models), whereas in the participant-rated models the psychological predictor varied depending on the functional domain being estimated. Verbal memory was the most common cognitive predictor for the clinician-rated measure, but measures of verbal fluency were more common for the participant-rated measure. Of the emotion-processing predictors, the pattern was less clear. Recognition of happy at low intensity was a significant predictor in 2 out of 5 models of the clinician-rated scale and none of the models of the participant-rated scale, whereas the exact opposite was true for recognition of disgust. Indices from the semantic stroop test were significant predictors in several of the models of participant-rated functioning, but only one of the models of clinician-rated function. However, it was not the same index in each model.

All in all there were not any striking differences between the estimations of the two different perspectives of functioning. It must be noted at this point that how many times a variable appears throughout the series of models is not necessarily an indicator that it is of broad or general importance. The dependent variables, especially those drawn from the same measure, are interrelated (see correlation matrix in Appendix 3, Table A6.54, on page 322) and therefore predictor variables related to one subscale of a functioning measure are likely to be related to other subscales of the same measure.

PATIENTS - DIFFERENCES BETWEEN FUNCTIONAL DOMAINS

Despite the interrelationship between the functional domains, the results in Table 9.43 show that work, interpersonal, and recreational domains are associated with different specific predictors. Work functioning showed the highest R^2, indicating it was the domain for which most variance was predicted of the three. Dysfunctional attitudes and self-esteem were the predictors that had the most explanatory power for clinician-

and self-rated work functioning respectively. Of the cognitive measures that reached significance in the models of work functioning, verbal memory, executive function measures and an index of a visual memory task (that was conjectured in Chapter 5 on page 122 to reflect impulsive responding and therefore potentially an executive failure) were significant predictors. Facial expression recognition (of disgusted faces and of happy faces) as well as a semantic stroop index were significant predictors from the emotion-processing variables. In the interpersonal domain, trait anxiety, verbal memory and executive function were significant predictors. In the recreational domain, dysfunctional attitudes and trait anxiety were significant predictors. In addition there was an association with one cognitive variable (category fluency), and with facial expression recognition and measures of the semantic stroop paradigm.

A pattern is difficult to discern. Generally each of the three domains showed a significant relationship with some psychological, some cognitive and some emotion-processing variables, but the specific indices differed. Work function was the only domain to show a significant relationship with self-esteem, whereas level of dysfunctional attitudes related significantly to both work functioning and recreational function. Trait anxiety significantly predicted both interpersonal and recreational function. Both interpersonal and work function were significantly predicted by verbal memory and executive measures. However, of the cognitive indices, recreational function was significantly associated only with an executive measure. Facial expression recognition was a significant predictor for both work and recreational function, but did not significantly predict interpersonal function.

Mood symptoms were only a significant predictor of total functioning (both self- and clinician-rated) but they did not significantly predict functioning in any of the three specific domains.

PATIENTS - THE ROLE OF DEMOGRAPHIC FACTORS AND CLINICAL VARIABLES

Adding demographic factors and clinical variables to the variables available to the regression models left the results almost unchanged. None of the demographic factors entered any model, and only age at onset entered the model predicting work function as measured by the LIFE-RIFT (β = -0.37, t_{32}=-3.18, p=0.003, ΔR^2=5.3, $F_{1,32}$=5.28, p=0.028). The lack of relationship with clinical variables is unsurprising as in the present sample there were almost no significant correlations between the functioning measures and illness history (see Appendix 6, Table A6.53, on page 321).

Table 9.43: Results of regression analyses in the patient sample with functioning as the dependent variable and cognitive, emotion-processing or psychological predictors as the independent variables. Predictors within each model are listed in the order in which they entered the model. β values are standardised coefficients and the R^2 reported is adjusted R^2.

Dependent	Independent	β	$t_{38\text{-}n}$	p	ΔR^2	$F_{n,39\text{-}n^*}$	p	R^2 %
Work								
LIFE-RIFT Work	DAS_{TOTAL}	0.40	3.60	0.001	14.1			
	$WMS_{LP\text{-}TOTAL}$	-0.47	-4.02	<0.001	7.7			
	$SDMTS_{0EASY}$	0.44	3.82	0.001	13.8	9.87	<0.001	53.9
	$Happy_{20}$	0.35	3.05	0.004	9.0			
	$Stroop_{ERRCW}$	0.31	2.80	0.008	9.3			
SASSR Work	SelfEsteem	-0.63	-5.93	<0.001	35.2			
	$SemStroop_{RTNEG}$	0.36	3.36	0.002	10.0	14.19	<0.001	58.1
	$VerFlu_{RULEBREAKS}$	0.29	2.70	0.011	7.6			
	$Disgust_{60}$	-0.25	-2.34	0.025	5.3			
Interpersonal								
LIFE-RIFT IP	TraitAnx	0.46	3.70	0.001	15.3			
	$WMS_{LP\text{-}TOTAL}$	-0.36	-2.64	0.012	14.3	7.82	<0.001	41.8
	$SDMTS_{0EASY}$	0.31	2.45	0.020	6.8			
	$CatFlu_{TOTAL}$	-0.28	-2.05	0.048	5.4			
Recreation								
LIFE-RIFT	DAS_{TOTAL}	0.52	4.40	<0.001	30.4			
	$Facial_{TOTAL}$	-0.45	-3.64	0.001	10.1	13.18	<0.001	49.0
	$SemStroop_{RTNEG}$	-0.33	-2.64	0.012	8.5			
SAS-SR	$CatFlu_{TOTAL}$	-0.41	-2.97	0.005	17.7			
	TraitAnx	0.32	2.37	0.024	8.4	7.25	0.001	33.0
	$SemStroop_{ERRPOS}$	0.30	2.18	0.036	6.9			

Table 9.43 continued

Dependent	Independent	β	t_{38-n}	p	ΔR^2	$F_{n,39-n*}$	p	R^2 %
Total								
GAF	Depression	-0.44	-3.75	0.001	19			
	$WMS_{LP-TOTAL}$	0.43	3.51	0.001	8.3			
	$SDMTS_{0EASY}$	-0.46	-3.45	0.002	6.9	8.04	<0.001	48.1
	$HSCT_B$	0.36	2.98	0.005	8.8			
	TMT_A	-0.27	-2.07	0.046	5.1			
LIFE-RIFT	DAS_{TOTAL}	0.62	6.61	<0.001	33.2			
	$WMS_{LP-TOTAL}$	-0.53	-5.42	<0.001	16.6			
	$SDMTS_{0EASY}$	0.32	3.31	0.002	7.1	16.84	<0.001	67.6
	$Stroop_{ERRCW}$	0.27	2.92	0.006	5.8			
	$Happy_{20}$	0.24	2.47	0.019	4.9			
SAS-SR	Depression	0.63	5.17	<0.001	16.6			
	$CatFlu_{TOTAL}$	-0.43	-3.85	0.001	12.8			
	$RAVLT_{A1}$	0.52	3.98	<0.001	10.5	12.35	<0.001	59.9
	$SemStroop_{ERRPOS}$	0.36	3.34	0.002	10.4			
	$Disgust_{60}$	-0.32	3.03	0.005	9.6			

*n = number of significant predictors

CONTROLS - GENERAL

It can be seen from Table 9.44 on page 254 that, firstly, fewer significant predictors entered the models of functioning in the control sample (in one case – LIFE-RIFT work – there were no significant predictors at all that entered the model). In general the control sample exhibited less variability in functioning than the patient group, showing ceiling effects on some of the measures (especially the clinician-rated measures). Explanatory power was similar for both groups, although lower for the control group than the patient group on most measures. It was notably lower in some

instances (e.g. clinician-rated total function on the LIFE-RIFT where 22.4% of the variance was explained in the control sample contrasted with 67.6% in the patient group). The main explanatory factors were from the psychological variables, especially trait anxiety, and cognitive variables were much less prevalent.

CONTROLS - CLINICAN-RATED VS SELF-RATED MEASURES

The comparison between clinician- and self-rated measures can only be made for recreational and total functioning. There were marked differences between the predictors of the two different types of ratings. The clinician-rated measures were significantly associated mostly with cognitive measures, whereas the self-rated measures were associated with psychological variables (self-esteem and trait anxiety specifically).

CONTROLS - DIFFERENCES BETWEEN FUNCTIONAL DOMAINS

For the control participants, work functioning was predicted by psychological factors (symptoms of depression and level of trait anxiety), interpersonal functioning was predicted by trait anxiety, verbal memory, and facial emotion recognition, whereas recreational functioning was predicted by cognitive measures and self-esteem. The largest proportion of variance in total functioning was explained by psychological factors (predominately trait anxiety), and some additional variance in clinician-rated total functioning from the GAF was explained by psychomotor and executive function measures.

CONTROLS - THE ROLE OF DEMOGRAPHIC FACTORS

Adding the demographic variables to the variables available to the models made almost no difference to the results. Age became a significant predictor in the model estimating interpersonal function (β=0.46, t_{22}=4.14, $p<0.001$) and years of formal education became a significant predictor in total self-rated function (β=0.41, t_{26}=4.02,

p=0.001). In both of these models, the addition of the demographic variable led to the addition of at least one further measure (interpersonal function: age and engagement with sad faces on the dot probe task replaced total facial recognition; total self-rated function: errors on the facial dot probe task and engagement with sad faces on the dot probe task also entered the model). The entry of other variables alongside these demographic measures may be indicative of multi-collinearity in these models (Studenmund, 2001). The variance proportion matrix on the original model of interpersonal functioning indicated potential problems with multi-collinearity (although there was no marked elevation of the variance inflation factor), which may explain the entry of extra variables alongside age in this model.

Table 9.44: Results of regression analyses in the control sample with functioning as the dependent variable and cognitive, emotion-processing or psychological predictors as the independent variables. Predictors within each model are listed in descending order of proportion variance explained. β values are standardised coefficients and the R^2 reported is adjusted R^2.

Dependent	Independent	β	t_{27-n}	p	ΔR^2	$F_{n,28-n*}$	p	R^2 %
Work								
LIFE-RIFT Work	None entered	-	-	-	-	-	-	-
SASSR Work	Depression	0.43	2.75	0.011	29.7	11.025	<0.001	42.6
	TraitAnx	0.41	2.62	0.015	12.9			
Interpersonal								
LIFE-RIFT IP	TraitAnx	0.53	3.70	0.001	18.6	8.44	<0.001	52.4
	$RAVLT_{A1}$	0.46	3.29	0.003	16.9			
	$Facial_{TOTAL}$	-0.31	-2.20	0.038	10.2			
	$DotProbe_{ERROR}$	-0.29	-2.10	0.047	6.7			
Recreation								
LIFE-RIFT	$Stroop_{ERRCW}$	0.32	1.96	0.062	33.3	10.02	<0.001	50.1
	$SDMTS_{0EASY}$	-0.35	-2.32	0.029	9.3			
	$DotProbe_{ERROR}$	0.33	2.17	0.040	7.5			
SAS-SR	SelfEsteem	-0.50	-2.89	0.008	21.4	8.37	0.008	21.4
Total								
GAF	Mania	-0.46	-3.37	0.003	22.5	10.35	<0.001	50.9
	$DSST_2$	0.51	3.74	0.001	19.3			
	$WCST_{TRIALS}$	0.33	2.38	0.025	9.1			
LIFE-RIFT	TraitAnx	0.50	2.96	0.006	22.4	8.79	0.006	22.4
SAS-SR	TraitAnx	0.76	5.93	<0.001	55.8	35.11	<0.001	55.8

*n = number of significant predictors

SUMMARY

Broadly speaking across both of the groups psychological predictors emerged as the variables predicting the largest amount of variance in functioning, irrespective of functional domain. Of the cognitive variables, measures of executive function and verbal memory were the most prominent predictors. There were several associations noted with both explicit and implicit emotion-processing variables. The implicit processing tasks capture both aspects of cognitive function and of emotion processing. Therefore it is not clear from the reported association between function and implicit emotion processing tasks whether the cognitive or emotion-processing component drove this association. Explanatory power was good across all models, although tended to be lower for the recreational and interpersonal domains than for work and total functioning. In the patient group, the pattern of variables indicated similar types of variables predicted both clinician- and self-rated functioning, although the specific variables were different across the models. The functional domains were predicted by different variables, although there was no clear or face-valid pattern to the results. There was no notable relationship between illness history factors and functioning in the patient group. In general, of the psychological factors, symptom levels were not the major predictors (apart from in the models of total functioning). Trait anxiety, dysfunctional attitudes and self-esteem generally explained a greater proportion of variance than symptom levels.

FACTOR ANALYSIS

COGNITION

Factor analysis of the cognitive variables produced a four factor solution (see Table 9.45). In total the four factors accounted jointly for 71.5% of the variance. Three of

the entered variables did not appear in the final model due to either a low factor loading or loading evenly onto more than one factor. The identified factors seem to involve a predominately executive factor (factor 1, which accounts for 22.6% of the total variance), a memory factor (factor 2, 20.1%), a speeded processing factor (factor 3, 17.0%) and the final factor is loaded onto by only a single test – the Stroop colour word test errors (factor 4, 11.8%). As this final factor contains only a single test index of a single test it is difficult to ascertain what underlying process it may reflect. For that reason the factor is named after the test (Stroop colour word, or Stroop CW), rather than named speculatively after a function or process purportedly captured by this measure.

Table 9.45: Results of factor analysis for the cognitive variables; factor loadings from the rotated solution.

	Factor			
	Executive	**Memory**	**Speeded Processing**	**Stroop CW**
	1	**2**	**3**	**4**
	22.6%	20.1%	17.0%	11.8%
Variable				
SDMTS (Total correct)	-0.796			
SOPT (Total errors)	0.785			
Wisconsin Card Sorting Test (Total errors)	0.697			
BADS Zoo Map Version 1 raw score	-0.646			
RAVLT (Long-delay recall)		0.913		
RAVLT (Total recall trials 1-5)		0.882		
DSST (Attempt 1)			-0.767	
Hayling Sentence Completion Test (B errors)			0.746	
TMTa (Time)			0.669	
Stroop Colour-Word (Errors)				0.901
Verbal fluency (Total correct)[a,b]	-	-	-	-
Category fluency (Total correct)[c]	-	-	-	-
WMS Logical Passages (Story units)[c]	-	-	-	-

[a] Excluded for low factor loading
[b] Excluded for low communality
[c] Excluded for loading evenly on more than one factor

EMOTION-PROCESSING

The factor analysis of the emotion-processing variables arrived at a four factor solution (see Table 9.46). Together the factors accounted for 71% of the total variance. The variables tended to load up according to the tests from which they were derived, with factor 1 representing the disengagement scores from the dot probe task (Disengagement, which accounted for 25.4% of the variance), factor 2 representing the engagement scores from the dot probe task (Engagement, 21.3%), factor 3 including both of the emotion recognition tasks (Emotion Identification, 12.7%) and the final factor including both of the interference indices derived from the Semantic Stroop paradigm (Semantic Stroop, 11.6%). In the initial analysis, total recognition for each of the five emotions of the facial expression recognition task were included as variables. However, analysis of the correlation matrix showed that these measures correlated poorly with other measures and three out of five showed poor sampling adequacy with KMO<0.5. They were therefore replaced with total recognition across all emotions in the final analysis.

Table 9.46: Results of the factor analysis of emotion-processing variables; factor loadings for the rotated solution.

	Factor			
	Disengagement	**Engagement**	**Emotion Identification**	**Semantic Stroop**
	1	**2**	**3**	**4**
	25.4%	21.3%	12.7%	11.6%
Variable				
Dot Probe - Disengagement Sad	0.908			
Dot Probe - Disengagement Disgusted	0.870			
Dot Probe - Disengagement Happy	0.844			
Dot Probe - Disengagement Angry	0.808			
Dot Probe - Engagement Disgusted		0.825		
Dot Probe - Engagement Sad		0.805		
Dot Probe - Engagement Happy		0.758		
Dot Probe - Engagement Angry		0.726		
Facial Expression Recognition (Total Correct)			0.872	
Vocal Emotion Recognition (Total correct)			0.813	
Semantic Stroop (Positive interference)				0.829
Semantic Stroop (Negative interference)				0.800
Recognition of individual emotions[a]	-	-	-	-

[a] Excluded for lack of correlation with other variables; replaced with total recognition

PSYCHOLOGICAL MEASURES

Factor analysis of the mood and other psychological measures produced a single factor (see Table 9.47). One variable was excluded from the analysis as it did not correlate with any of the other measures. Four variables were left, each with a strong loading onto a single factor, which accounted for 71.2% of the variance. This factor was labelled 'Psychological', as it incorporates scores on measures of psychological constructs such as self-esteem and dysfunctional attitudes, alongside mood and anxiety.

Table 9.47: Results of the factor analysis of psychological variables; factor loadings of the rotated solution

	Factor
	Psychological
	1
	71.2%
Variable	
Trait Anxiety Inventory score	0.901
Rosenberg Self Esteem Scale score	-0.890
Beck Depression Inventory score	0.805
Dysfunctional Attitudes Scale total score	0.772
Altman Mania Rating Scale score[a]	-

[a] Excluded for lack of correlation with other variables and KMO<0.5

FINAL OVERALL MODEL OF FUNCTIONING

The results of the regression analyses using the factor scores derived from the factor analyses as the independent variables are reported separately for the patients (Table 9.48, page 261) and controls (Table 9.49, page 262). The main finding is that the scores on the psychological factor explained the largest proportion of variance in almost all of the models. In the control sample, the factor scores on the psychological factor were the only variable to enter 5 out of 8 of the models. Of the remaining three, clinician-rated work function was significantly associated with scores on the speeded processing factor, clinician-rated total function (with the GAF) was significantly associated with scores on the stroop colour word factor, and no variables entered the model predicting clinician-rated recreational function.

For the patient group, the psychological factor appeared in every model and was the largest predictor in all but one in terms of variance explained. It was the only significant predictor in three of the models. The psychological and executive factors

together were the only predictors in a further two models (where both variables explained a very similar proportion of variance). The final three contained the psychological factor alongside either the Stroop colour-word factor, the emotion identification factor, or both the Stroop colour-word factor and the semantic Stroop factor. In the patient group, work function was predicted by a combination of the psychological factor and the Stroop colour-word test factor, interpersonal function was predicted only by the psychological factor, and recreation was predicted by a combination of the psychological factor, the executive factor and the emotional identification factor. Total functioning was predicted mostly by the psychological factor and one of two cognitive factors (Stroop colour word or executive).

Across all models explanatory power was notably lower in both groups when contrasted with the previous models (patient group range: 14.5%-44.9%, control group range: 10.7%-49.5%).

Table 9.48: Results of regression analyses in the patient sample with functioning as the dependent variable and cognitive factor scores, emotion processing factor scores and psychological factor scores as the independent variables. Predictors within each model are listed in descending order of proportion variance explained. β values are standardised coefficients and the R^2 reported is adjusted R^2.

Dependent	Independent	β	t_{38-n}	p	ΔR^2	$F_{n,39-n*}$	p	R^2 %
Work								
LIFE-RIFT	Psychological	0.41	2.91	0.006	13.8	7.50	0.002	25.5
	Stroop CW	0.37	2.61	0.013	11.7			
SASSR	Psychological	0.50	3.50	0.001	22.8	12.24	0.001	22.8
Interpersonal								
LIFE-RIFT	Psychological	0.41	2.72	0.010	14.5	7.42	0.010	14.5
Recreation								
LIFE-RIFT	Psychological	0.43	3.07	0.004	17.3	7.71	0.002	26.1
	Emotion ID	-0.33	-2.33	0.026	8.8			
SAS-SR	Executive	0.38	2.58	0.014	10.6	5.59	0.008	19.5
	Psychological	0.33	2.25	0.031	8.9			
Total								
GAF	Psychological	-0.50	-3.48	0.001	22.6	12.07	0.001	22.6
LIFE-RIFT	Psychological	0.63	5.09	<0.001	29.1	11.32	<0.001	44.9
	Stroop CW	0.31	2.55	0.015	9.5			
	Semantic Str	-0.28	-2.26	0.03	6.3			
SAS-SR	Psychological	0.45	3.35	0.002	16.5	9.47	<0.001	30.8
	Executive	0.40	2.94	0.006	14.3			

*n = number of significant predictors

[a] Semantic Stroop

Table 9.49: Results of regression analyses in the control sample with functioning as the dependent variable and cognitive factor scores, emotion processing factor scores and psychological factor scores as the independent variables. Predictors within each model are listed in descending order of proportion variance explained. β values are standardised coefficients and the R^2 reported is adjusted R^2.

Dependent	Independent	β	t_{27-n}	p	ΔR^2	$F_{n,28-n*}$	p	R^2 %
Work								
LIFE-RIFT	Speeded Processing	0.38	2.09	0.046	11.1	4.37	0.046	11.1
SASSR	Psychological	0.59	3.75	0.001	32.6	14.09	0.001	32.6
Interpersonal								
LIFE-RIFT	Psychological	0.50	2.92	0.007	21.8	8.54	0.007	21.8
Recreation								
LIFE-RIFT	None entered	-	-	-	-	-	-	-
SAS-SR	Psychological	0.42	2.33	0.028	14.1	5.45	0.028	14.1
Total								
GAF	Stroop CW	0.37	2.06	0.050	10.7	4.24	0.050	10.7
LIFE-RIFT	Psychological	0.45	2.59	0.016	17.4	6.7	0.016	17.4
SAS-SR	Psychological	0.72	5.25	<0.001	49.5	27.52	<0.001	49.5

LEARNING POTENTIAL AND FUNCTIONING

The correlations between the measures of learning potential and functioning are shown in Table 9.50 below. There was a significant positive relationship in the patient group for percentage change between the last two administrations of the Digit Symbol Substitution Test (trials 3 and 4) and clinician-rated work function, self-rated work function, self-rated recreational functioning, and clinician-rated total function. The correlations were all positive, indicating that a higher percentage change in performance between the final two trials was associated with a poorer level of functioning. In the control group, the same relationship was not evident. There was only a single statistically significant negative correlation between self-rated recreational function and percentage change between trials 2 and 3.

Table 9.50: Spearman correlations between learning potential measures derived from the Digit Symbol Substitution Test (DSST) and functioning measures

	GAF	LIFE-RIFT				SAS-SR		
		Work	Interpersonal	Recreation	Total	Work	Social & Leisure	Total
Controls								
DSST %Δ trial 1 to trial 2	0.36	0.15	-0.19	-0.17	-0.08	-0.22	-0.07	-0.07
DSST %Δ trial 2 to trial 3	0.03	0.30	0.02	0.10	0.21	-0.20	-0.52*	-0.13
DSST %Δ trial 3 to trial 4	-0.17	0.04	0.07	0.19	0.12	0.35‡	0.16	0.18
Slope	0.25	0.38‡	-0.23	0.11	0.10	-0.16	-0.37‡	-0.07
Bipolar Patients								
DSST %Δ trial 1 to trial 2	-0.07	-0.09	-0.03	-0.08	0.05	0.05	0.14	0.11
DSST %Δ trial 2 to trial 3	0.02	0.25	0.06	0.17	0.10	0.39*	0.16	-0.10
DSST %Δ trial 3 to trial 4	0.19	0.37*	0.18	0.26	0.39*	0.42*	0.39*	-0.11
Slope	0.02	0.14	0.13	0.05	0.19	0.30‡	0.27	-0.09

‡ 0.1<p<0.05, *p<0.05, ** p<0.01
DSST %Δ, Digit Symbol Substitution Test % change

DISCUSSION

SUMMARY – BETWEEN-GROUP RESULTS

Patients with bipolar disorder showed significant impairment in almost all aspects of functioning as assessed by both clinician- and self-rated measures. Work, interpersonal relationships, recreation, and role functioning – all as captured by a variety of different functional rating scales – showed significant and marked impairment. Although not compared directly, the degree of impairment as rated by patients themselves was of a similar magnitude to that rated by clinicians, with all measures showing some indices with large effect sizes (i.e. d>0.8). There were some domains in which patients did not show significant impairment, namely physical functioning (e.g. use of limbs & sense organs), personal care (washing, dressing and feeding onself), leisure activities with friends and family, and parental/family functioning. Some of these areas would not be expected to show impairment in this patient group (e.g. physical functioning and personal care), especially not in a euthymic sample. Others did not apply to the whole sample and may have suffered low power as a result (e.g. parental functioning).

SUMMARY - PREDICTING FUNCTIONING

In the first set of regression analyses, which used raw test or questionnaire scores as predictor variables, broadly speaking across both of the groups psychological predictors emerged as the variables predicting the largest amount of variance in functioning, irrespective of functional domain. The models in the patient group included a larger number of predictor variables and generally explained more variance than the models in control participants. Cognitive variables were significantly associated with

functioning in the patient group more so than in the control group. Of the cognitive variables that were significant predictors of functioning in the patient group, measures of executive function and verbal memory explained the most variance. There were several associations noted with both explicit and implicit emotion-processing variables, although no consistent pattern was evident. Explanatory power was good across all models, although tended to be lower for the recreational and interpersonal domains than for work and total functioning. In the patient group, the pattern of variables indicated similar types of variables predicted both clinician- and self-rated functioning (psychological predictors were primary in terms of proportion variance explained, followed by cognitive predictors, then emotion-processing predictors), although the specific variables were different across the models. The different functional domains (work, interpersonal and recreation) were predicted by different variables, but in the patient group there was not a clear or face-valid pattern (such as cognitive variables predicting work function, emotion-processing variables predicting interpersonal function). There was no notable relationship between illness history factors and functioning in the patient group. Current level of symptoms was generally not a significant predictor (apart from in the models of total functioning). Of the psychological variables, trait anxiety, dysfunctional attitudes and self-esteem generally explained a greater proportion of variance than symptom levels.

The second stage of the analyses used a data reduction technique to reduce the potential number of explanatory variables. The factor structure of the data indicated that several strong and coherent factors could be derived and in general the factors had a straightforward interpretation. The emotion-processing variables loaded up according to the task from which they were derived, indicating the tasks did not tap similar processes. In the multiple regression analyses using factor scores as the independent variables, the psychological factor explained the largest proportion of variance in almost

all models in both the patient and control groups. In the control group, the psychological factor was the only variable to enter the model in five out of the eight equations. For the patient group, a very similar pattern emerged. The psychological factor predicted the largest proportion of functioning in almost all of the models. The executive factor and Stroop colour word factor were the next most important predictors in terms of proportion of variance explained. In both groups, the proportion of variance explained by the factor scores was notably less than that explained by the raw scores.

Finally, the correlations between the learning potential measures and functioning measures indicated that a larger percentage gain between the final two administrations of the Digit Symbol Substitution Test was associated with worse functioning in the patient group. The implications of this finding are discussed in more detail below.

INTERPRETATION AND RELATIONSHIP WITH PREVIOUS FINDINGS

The current findings extend previous work investigating the relationship between cognitive function and social function in this patient group. Studies using regression analyses to explore cognitive predictors of functioning in bipolar patients have reported mixed findings, with some reporting a preferential association with cognitive variables (Martinez-Aran et al., 2007), some with clinical factors (Laes and Sponheim, 2006), and some with both (Dickerson et al., 2004, Simonsen et al., 2010). The proportion of variance explained has tended to be low to moderate (8%-50%). In the present study, an association was reported between functioning measures from three domains – psychological functioning, cognitive function and emotion-processing – and the proportion of variance explained by the models using raw scores as predictors was generally better than that seen in previous studies (ranging from 33% to 68% in the patient group).

The inclusion of a broader selection of measures of psychological functioning, including dysfunctional attitudes, trait anxiety and self-esteem, seems to have been the major factor in increasing explanatory power. There have been few investigations of the relationship between these constructs (especially dysfunctional attitudes and self-esteem) and social functioning in bipolar disorder to date. One investigation reported no correlation between self-esteem and social adjustment in a mixed sample of patients with either unipolar major depression or bipolar disorder (Serretti et al., 1999). A different study in a mixed psychiatric sample reported that social functioning of individuals involved in a service user run self-help service was significantly correlated with what the authors termed psychological factors – level of hope, self-efficacy and use of problem-centred coping mechanisms (Yanos et al., 2001). Given the relationships demonstrated in the present investigation, further exploration of psychological factors in bipolar disorder and their relationship with social functioning is necessary.

One study identified marked instability in the self-esteem of bipolar patients over the course of a single week (Knowles et al., 2007). This may be of significant relevance for functioning. Self-esteem was identified as a significant predictor of self-rated work functioning in the current study, and if future investigations can demonstrate that, in this context of exploring predictors of social functioning, self-esteem is operating as more than simply a broad subjective measure of how an individual gauges themselves to be functioning, it may be that addressing poor or unstable self-esteem could have a positive impact on work functioning. The instability of self-esteem in patients with bipolar disorder also raises issues about when to measure it and whether the relationship between self-esteem and functioning will vary according to the patient's mood state. It was identified in chapter 3 that different ways of assessing self-esteem (explicit versus implicit methods) can produce opposite results in the same person, especially during periods of manic symptoms. Self-esteem measured explicitly has been

shown to be more positive than when measured implicitly in manic patients (Lyon et al., 1999), so the relationship between self-esteem and function may well be complex and a fruitful source of further investigation.

The relationships with cognitive variables demonstrated here in the regression analyses replicate previous findings. Previous studies have reported associations between social functioning and executive function (Atre-Vaidya et al., 1998, Martinez-Aran et al., 2004a, Martinez-Aran et al., 2004b, Zubieta et al., 2001) and between social functioning and verbal memory (Atre-Vaidya et al., 1998, Dickerson et al., 2004, Martinez-Aran et al., 2004a, Martinez-Aran et al., 2004b, Martinez-Aran et al., 2007). In general in the present results, the verbal memory measure predicted clinician-rated functioning whereas various aspects of executive function predicted self-rated function. This may explain why the one other study in a euthymic patient sample that used regression analyses to investigate cognitive and clinical predictors of functioning reported a relationship with verbal memory and not with executive function – the functional measure used was the clinician-rated GAF (Martinez-Aran et al., 2007). It remains to be understood why it should be the case that memory function predicts a larger proportion of variance in clinician-rated function than executive measures and vice-versa for self-rated measures. Clinican-rated measures often rely on verbal reports from patients so it may be there is a general verbal factor involved that links the two. More studies using both types of ratings are needed to see if the current findings are replicated before any conclusions can be drawn.

There was no significant relationship demonstrated between illness history variables and social function in the present sample. Again, this is consistent with another similar study in a euthymic sample (Martinez-Aran et al., 2007). Additionally, current symptoms of depression and mania were not significant predictors of social functioning

in patients, apart from measures of total functioning, which were significantly associated with self-rated depressive symptoms. Studies that have investigated both cognitive and clinical factors using regression analyses have tended to show a significant relationship with symptoms over and above cognition only in non-euthymic samples (Laes and Sponheim, 2006, Simonsen et al., 2010). Given that functioning is – by definition – impaired in a mood episode, it seems likely that during the episode the severity of mood symptoms will be the key determinant of functioning. Only between episodes does the relationship with other factors become more evident. It must be noted, however, that measures such as dysfunctional attitudes and self-esteem, two of the psychological factors that were major predictors of social functioning in this euthymic sample, may be operating as broad measures of residual symptoms. These measures may capture more effectively the disabling aspects of inter-episode symptoms compared to standard symptom measures of depression and mania.

The analyses indicated that, by and large, both self- and clinician-rated functioning are predicted by very similar measures. Whilst the first stage analyses indicated differences in terms of the specific individual variables that entered different models, when replaced by factors, these differences were less evident. A similar picture was true of the predictors of functioning in different domains. The analyses incorporating raw scores identified different predictors for the different functioning domains, but the pattern was more similar when considering the analyses using factor scores. In the patient group, both work and recreation were predicted by the psychological factor and a cognitive factor, although for the work domain this was the Stroop Colour-Word factor (potentially indexing inhibition) whereas for the recreation domain this was the broader executive factor. Emotion identification was a significant predictor of the recreation domain only. These small differences may be important, or the broad similarity may indicate that the factor analyses produced an outcome that was

too crude to identify subtle differences in the aspects that impact on different functional domains. It may also be that, fundamentally, the core processes that impact on different areas of functioning are highly similar, but artifacts of measurement create the appearance that individual indices are highly specific predictors of particular functional domains. One difficulty in the present data set that may also have created a convergence in the predictors of the different functional domains was the fact that the divisions of functional categories were more apparent than real, that is scores on the different domains were highly correlated with one another. This could explain why similar predictors appeared in different regression models. However, to date this is the first study that has investigated cognitive predictors of different functional domains using both self- and clinician-rated measures. It is necessary to replicate the present findings before drawing conclusions about which predictors relate most closely to different functional domains.

One further aspect worthy of note is the relatively weak relationship between social function and emotion processing variables in this sample. A relationship between emotion identification and both work functioning and independent living has been noted in patients with schizophrenia (Kee et al., 2003). This finding was not replicated in this sample of patients with bipolar disorder. Chapter 8 demonstrated that there was no evidence of major impairment in any aspect of emotion-processing in the present sample. If further research establishes that emotion-processing is relatively unaffected in bipolar disorder, it may explain the absence of association with social functioning and also represent an important divergence between schizophrenia and bipolar disorder.

The final result to discuss from the present study is the association between learning potential and social function. Green (1996) proposed that learning potential – a notion of latent capacity, or of the amount of information an individual is capable of

learning – may be the key cognitive skill that underlies the association between cognitive function, skills acquisition and social functioning in patients with Schizophrenia. In the present sample, repeat performance on the Digit Symbol Substitution Test was used as a measure of learning potential and percentage change between each administration as well as overall learning slope were correlated with social functioning. There was a negative relationship between percentage gain between the final and penultimate trials and several of the indices of social function. A greater gain between the final two administrations was associated with poorer functioning. This may reflect that those who continue to gain on later administrations were 'slower learners' on earlier trials. It is possible that individuals with better functioning gain rapidly in the first two or three administrations and then plateau (in Chapter 5 on page 126 the control participants showed this pattern and a statistically significant difference between the two groups was only evident for percentage change between trials one and two), whereas those with a smaller initial gain continue to learn more slowly. Further exploration of this finding to establish whether this pattern is evident is necessary before conclusions can be drawn. Although the relationship was relatively weak (significant correlations were in the region of 0.4), the result nonetheless provides some evidence that Green's hypothesis may also relate to patients with bipolar disorder.

LIMITATIONS

There are number of limitations of the present study that it is necessary to consider.

The first limitations pertain to the statistical analysis. Both sets of analyses predicting function suffer drawbacks. The major concerns in the first stage analyses using test indices as predictor variables in a forwards-entry stepwise procedure are the lack of a hypothesis-driven approach to variable entry and multicollinearity. A data-driven

approach was selected in the absence of specific hypotheses. This can lead to surprising or unexpected findings (although this was not the case). One of the major difficulties in the present study was identifying potential predictor variables from the large number of possibilities. An alternative approach to the one taken could have been to generate specific hypotheses on the basis of previous findings in bipolar disorder or Schizophrenia and derive composite measures from standardised scores to reduce the number of predictors. The difficulty with this approach is it often reduces comparability between studies. However, future studies should consider using this approach, especially as the available literature grows and there is an increasingly-solid foundation of data to use to derive hypotheses. With regard to multi-collinearity, there was a large number of potential independent variables selected as possible predictors and they are likely to show some degree of interrelatedness, in part because the various tasks used draw upon some common cognitive resources and also because one individual provides multiple data points. In the case where two variables are highly related and both explain the same portion of variance in the dependent variable, it is not possible for the statistical model to identify the 'true' predictor. The result is that regression equations estimated on highly collinear data can be unstable (e.g. the removal of one non-significant variable can lead to a large change in the model as a whole). It remains possible that variables that did not enter the equation could replace variables that did and perform an almost-equivalent job (Studenmund, 2001). Additionally the estimates of the coefficients are less precise which can affect their statistical significance (which in a stepwise entry procedure can also affect the composition of the final model) (Studenmund, 2001). Whilst the covariance matrices and variance inflation factors indicated very few problems due to multi-collinearity, in several models the sign on the coefficients were the opposite to that expected. This is a phenomenon that can be caused by multi-collinearity (Studenmund, 2001), suggesting this was a problem for some of the models.

However, regression equations estimated on data showing multi-collinearity are still statistically valid (that is, the estimated coefficients are still best linear unbiased estimates), the overall fit of the equation is not affected, and generalisability (or forecasting) is not affected provided the same collinearity exists in the data set for which the model is used to make predictions (Studenmund, 2001). With regard to this latter point, it was not the aim of the regression analyses in this study to produce an equation that explained functioning in bipolar disorder that could be used as the basis to make predictions in different samples. It is therefore likely that the major impact of multi-collinearity on the present data is on the composition of the final models. The main value of this approach was that it provided results that were more easily compared to previous studies in bipolar disorder, which to date have all used raw scores rather than reduced data or composite measures.

The factor analysis that was used to minimise problems resulting from collinearity also has limitations. The data-reduction technique used (principal components analysis) produces results that are grounded in this data set. This may limit the extent to which the results generalise. From one perspective, it could also be argued that a new level of abstraction has been added into the data by using principal components rather than raw data scores to relate to functioning. This may detract from a clear, real-world interpretation of the data. However, in the present case at least, it is not clear that this was the case. The variables loaded onto the components in a readily-interpretable way and the regression analyses produced results that were consistent with the preceding set of results. However, it cannot not be ruled out that the factors were over-general and failed to capture some of the important variance that the individual components brought to the prediction of social function. This may explain why similar factors appeared in each regression model and why the variance explained was

lower in the regression models using factor scores compared to the individual raw scores.

It is possible that, owing to the relatively small sample size, some of the models were overfitted (containing more variables than is reasonable for the number of observations). It is therefore necessary to replicate the present findings in a larger sample.

The domain scores on the functioning measures were used directly as dependent variables assuming that the validation of the relevant measures had resulted in true and independent measures of 'work functioning', 'interpersonal functioning', and 'recreational functioning'. However, as mentioned above, exploration of the correlation matrices between the various measures indicated a notable degree of overlapping variance. Using factor analysis to derive functioning factors may have minimised difficulties resulting from this covariance, perhaps clarifying the picture and reducing the number of models that were estimated. This is a valuable approach that should be taken forward in future investigations. To retain comparability with previous studies (especially those using the GAF), it was not employed here.

The study was cross-sectional and not longitudinal. The direction of causation cannot be determined. It is plausible, and indeed likely, that once the illness is established there is a reciprocal relationship between social functioning, psychological factors, and cognitive function, potentially establishing a vicious circle of worsening functional impairment and poorer clinical outcome. The direction of causation has important implications for intervention and which side of the equation to target in order to improve patient outcomes.

Although no relationship with illness history was reported in the present study, illness history data is very difficult to collect accurately and in the present study self-report was used. This data is therefore likely to be subject to bias and the results may not have accurately reflected the real impact that previous experience of illness has on current functioning.

Finally, inclusion of an objective function measure (such as occupation or earnings) could have been included to minimise subjective biases in functioning judgments. This may have added greater external validity to the findings (although what constitutes an objective measure can be a contentious issue for some areas of functioning).

CONCLUSION & FUTURE DIRECTIONS

The major finding of the present analyses is that social functioning is strongly predicted by psychological factors such as self-esteem, dysfunctional attitudes and trait anxiety. Cognitive function, particularly executive function and verbal memory, shows a significant association with functioning, but explains less variance than psychological factors. The implications for individuals with bipolar disorder are that efforts focused on improving psychological factors should be the primary targets in order to improve functional outcomes. However, the cross-sectional nature of this data does not indicate the direction of causation. A valuable next step would be investigating whether improving self-esteem does indeed improve functional outcomes in a longitudinal prospective study.

There are still a number of outstanding issues surrounding this area and also a number of methodological issues that should be addressed by future research. Turning to the methodological issues first, many previous studies have measured functioning close to the end of an episode. It is not clear how long is needed for full functioning to be

regained and some aspects of functioning are easier to regain (e.g. housing situation, friendships) than others (e.g. work, where there may be institutional barriers to ongoing employment). At the very least future research should take careful note of the symptom status and length of time since an episode for their participants. If it is not possible/feasible to standardise these indices, the information can be controlled for, or any relationships explored. A second major issue in the literature to date is the variety of ways in which analyses are conducted and reported. It is common practice to report only the relationships with functioning that reached statistical significance. In many instances it is left unclear which relationships have been assessed. This produces bias in reported results and restricts any reasonable attempt to amalgamate findings. Many of the samples used have been relatively small and non-significant results may still be of a reasonable magnitude. This difficulty is compounded by the variety of social functioning measures used further reducing comparability across studies. When correlations or regression analyses are used to investigate associations with social functioning measures, the full analysis strategy should be clearly outlined and the reporting criteria for the results stated.

With regard to some of the other remaining issues, one important question surrounds whether an individual's functional limitations are imposed upon them due to their illness, or whether in effect they represent an adaptive coping mechanism. This issue is particularly salient with regard to work functioning. Work functioning showed the largest (or second-largest in one case) effect size of all of other subscales (excluding 'total') on the four measures that included a clearly identifiable work subscale (LIFE-RIFT, SLOF, LFQ, & SAS-SR). However, have individuals chosen not to work in order to minimise exposure to stressful situations that may exacerbate their illness, or are their symptoms (or self-esteem) so debilitating as to render work an impossibility? Whilst recovery defined in socio-economic terms is clear and readily-understandable, it may oversimplify

matters. Returning to work may not be a highly valued goal for some people, who instead value stability of their mental health and strong relationships. It must be ascertained, perhaps through qualitative research, that improving an individual's ability to work would genuinely enhance their overall quality of life rather than simply score them more highly on a functional outcome measure. One qualitative study has shown that returning to work confers significant benefits for people with mental illness, giving a sense of meaning and promoting recovery (Dunn et al., 2008). However, this is not the case for everybody with mental illness and further exploration of the relationship between work, wellness and bipolar disorder would be beneficial to understand how important a personal goal it is for patients to be employed.

The final issue touches on something that was not explored in the present study – the environmental barriers to social functioning. Bipolar disorder is usually a chronic, episodic illness. An individual not only takes the 'blow' of the acute episode, but then also faces the aftermath. For example, impulsive spending during a manic episode can result in severe debt; promiscuity can damage long-term relationships; friends, colleagues and jobs can be lost due to inappropriate behavior or comments during periods of grandiosity or irritability. Social relationships can suffer greatly during periods of withdrawn behavior during a depressive episode. Patients who regularly have to reconstruct their life and face the stigma associated with mental illness may understandably find it difficult return to an earlier level of functioning time and time again. Acute episodes can result in long periods off work, perhaps resulting in gaps in work record and becoming deskilled. It may become difficult to find employment or prove too daunting a task to retrain. Although the legislative framework is designed to protect against disability discrimination, it can only be used if employees are able to be open with their employer about their mental health difficulties, something that fear of stigma often prevents.

There are therefore many understandable reasons why functioning is often impaired in people with bipolar disorder. Improving self-esteem and other psychological factors (and, to some extent, cognitive function) may go some of the way to remediating that, but that would only address one side of the equation. A comprehensive research agenda is needed that explores both the individual-level factors that impede functioning and also the societal barriers in order to develop a thorough understanding of where change or intervention is needed to improve patients' functional outcomes.

CHAPTER 10: DISCUSSION

SUMMARY OF FINDINGS

The aim of the present investigation was to further explore aspects of cognitive function, emotion-processing and social functioning in euthymic patients with bipolar disorder. There were a number of questions that were addressed, namely: is there a relationship between memory and executive impairment? Can cognitive dysfunction be improved by a self-monitoring intervention? Are deficits evident in social-information processing? And do cognitive dysfunction and emotion-processing impact significantly on psychosocial function? The theme connecting these various elements was to understand not only the nature and degree of cognitive impairment in bipolar disorder, but also the impact it has on patients' everyday lives.

This study has demonstrated that, consistent with a large number of previous studies, patients with bipolar disorder showed impaired performance on measures of verbal learning & memory, visual memory, psychomotor processing and multiple aspects of executive function (inhibition, planning, and verbal fluency).

Further exploration of the relationship between verbal memory impairment and executive dysfunction revealed that the degree of subjective organisation in verbal list-learning (a potential executive component of verbal memory) did not show deficits that were sufficient to explain the full extent of the memory impairment in this population. However, some questions remained about the indices of subjective organisation and which processes they measure.

In the next stage of the project, an intervention designed to promote self-monitoring while performing a neuropsychological task was used to explore whether this strategy would enhance performance. The results indicated that self-monitoring was associated with a smaller deficit on the Wisconsin Card Sorting Test in euthymic patients with bipolar disorder and the differences from control participants did not reach statistical significance. This is in contrast with the deficits shown by the patients who received the standard administration. On two indices, the patients who had engaged in self-monitoring significantly out-performed the patients who completed the standard administration. There is some evidence that self-monitoring assisted performance by enhancing concept formation and helping participants remember the current sorting principle.

The exploration of emotion-processing in general indicated no deficits in the patients with bipolar disorder. There were no differences between the groups in facial or vocal expression recognition. No differences in attentional engagement or disengagement with facial expressions of anger, disgust, happiness or sadness were noted in patients compared to controls. Patients showed greater interference from depression-related words compared to mania-related words on an emotional Stroop task suggestive of biased processing of depression-related stimuli.

In the final part of the study, the level of impairment in psychosocial function was investigated in patients using a number of clinician- and self-rated measures. Some of these measures were then used to explore cognitive, emotion-processing, and psychological predictors of social functioning. The results indicated that patients with bipolar disorder showed significant impairment in almost all aspects of functioning as assessed by both clinician- and self-rated measures. Work, interpersonal relationships, recreation, and role functioning – all as captured by a variety of different functional

rating scales – showed significant and marked impairment. There was no evidence of impairment in physical functioning, personal care, leisure activities with friends and family, or parental/family functioning.

With regard to predictors of functioning, psychological factors such as self-esteem, trait anxiety, dysfunctional attitudes and depressive symptoms explained a significant and large proportion of variance in social functioning. Measures of executive function and memory explained a relatively smaller amount of additional variance over and above that associated with the psychological factors. Measures of emotion-processing explained a very small proportion of variance in social functioning.

DISCUSSION

The results of the cognitive investigations indicated that there is limited support that executive dysfunction and memory impairment are related in euthymic patients with bipolar disorder. On the state of the current evidence it appears that the two are both distinct areas of dysfunction. This is in accord with correlational evidence suggesting the two broad areas relate differently to illness history variables, with executive dysfunction showing little relationship with illness progression but memory impairment showing a stronger association with poorer clinical outcome (Robinson and Ferrier, 2006).

Performance on at least some cognitive tasks could be markedly improved by a relatively straightforward self-monitoring intervention. This study is the first to use this intervention in patients with bipolar disorder and the promising result suggests that cognitive dysfunction in patients with bipolar disorder is modifiable, i.e. it is not a fixed, immutable deficit. The use of interventions that promote focus, enhance short-term

memory, and increase reasoning may be useful for patients in their everyday activities. This remains to be explored further in patients with bipolar disorder.

The results indicated no support for a marked impairment in emotion-processing in this group. Any effects are at best subtle. Patients made significantly more errors than controls on an attentional task incorporating emotional stimuli and an emotional Stroop task. However, there was little evidence of differences in response to the emotional stimuli indicating that these errors more likely stemmed from cognitive difficulties than interference caused by the emotional stimuli. The only exception was the finding that patients were slower to name depression-related words than neutral words on the emotional Stroop task. A similar bias has been noted in manic patients with bipolar disorder (Lyon et al., 1999). Although previous studies have interpreted this bias in terms of the depression avoidance hypothesis, it may instead indicate that there is an underlying sensitivity to depression-related stimuli that can be observed throughout all phases of the illness and may play some role in vulnerability to mood episodes. Whether this bias relates to factors such as low self-esteem and elevated dysfunctional attitudes that also remain in euthymia needs to be explored further in future studies.

The outcome of the exploratory investigation of predictors of social functioning indicated that efforts to improve cognitive function are unlikely to have a large impact on social functioning, as cognition accounted for only 5-10% additional variance in functioning once the impact of psychological factors had been taken into account. Instead, interventions designed to improve self-esteem, dysfunctional attitudes and trait anxiety would be better candidates for having a positive impact on functional outcome.

One apparent inconsistency deserves further discussion. There is an apparent mismatch between the degree of cognitive impairment in bipolar patients (in terms of effect size) and impact it has on social functioning. There is much emphasis placed on the

size of the cognitive deficit in this patient group, yet it does not have a correspondingly large impact on social functioning. One possibility is that formal assessments of cognitive function in bipolar disorder underestimate patients' abilities and in the real world the cognitive deficits are obviated or patients have developed coping strategies. It is possible that the tests being used lack ecological validity. Some of the differences noted in the patients group in the present investigation – for example lower self-esteem, higher dysfunctional attitudes and higher anxiety – may explain some divergence between the outcome of formal testing and an individual's actual level of ability. Neuropsychological testing can be daunting and anxiety-provoking, and individuals prone to think negatively about themselves may struggle when difficulties are encountered. Previous investigations rarely account for these variables and although psychological factors such as self-esteem and dysfunctional attitudes were measured in the present study, their relationship with neuropsychological test performance was not explored. Although there is a strong call for interventions to ameliorate cognitive deficits in this patient group, it is often made under the assumption that success will show via improved social function. It remains to be seen whether this is indeed the best route to improving function, and present evidence does not suggest that it is.

LIMITATIONS

The limitations within each area have been discussed at length in each chapter. The key limitations and those which affected the study as a whole are considered here.

There are a number of issues pertaining to the present sample. One limiting factor is the sample size. Whilst not out of keeping with similar studies, the sample is nonetheless relatively small. A larger sample would permit greater confidence in the findings and improved statistical power. Additionally, the sample was mostly recruited from tertiary care services thereby likely over-representing individuals with a more

severe course of illness. This may affect the extent to which the results will generalise to other patients. The criteria for euthymia required one month with very low symptom levels. Although many patients had been euthymic for longer than this, nonetheless questions remain whether this is sufficient time for recovery from an episode and a longer time period may reduce potential interference effects from ongoing symptomatic recovery. The illness history characteristics of the sample were derived from self-report. Conducting a more detailed structured interview or reviewing clinical notes may have made this data more reliable. Addtionally, all the patients were taking medication at the time of testing. The independent effects of medication on cognitive function were not assessed and it cannot be ruled out that treatment effects played a role in the deficits shown by the patients.

The study focused on executive function and verbal memory, and did not explore other aspects of cognitive function in depth (e.g. visual memory, spatial memory). It may be that these aspects of cognitive function are important for social functioning and should be explored in future studies.

FUTURE DIRECTIONS

In the process of addressing the aims of the present study, a number of other questions were raised and areas for further exploration highlighted.

Questions remain about the fundamental nature of cognitive impairment in this group. The present findings indicate that verbal memory impairment is not due to organisational difficulties in list-learning. There was a suggestion that retrieval-based learning may be impairmed in patients with bipolar disorder and this should be explored using tasks designed to assess retrieval-based learning.

In the emotion-processing domain, the implicit emotion tasks showed no or only subtle evidence of different processing of emotional stimuli in the patient group. This work should be extended using paradigms that have shown evidence of bias in patients with major depression, i.e. paradigms using longer stimulus display times that permit rumination and stimuli that are judged in relation to the self. These may prove more optimal conditions by which to assess emotion-processing biases in patients with bipolar disorder.

One question raised was whether the measures and techniques used to assess cognitive function in bipolar disorder accurately represent an individual's level of ability. The use of supportive interventions, such as self-monitoring, and increased ecological validity of tests should be explored further. Moving tests closer to situations involved in the real world may help clarify whether ecological validity is indeed an issue.

With regard to psychosocial function, very little emphasis was placed on identifying the strengths and weaknesses of the functioning measures employed here for capturing the relevant concepts within patients who have bipolar disorder. There has been much debate in the literature about measures appropriate for use in this patient group. The strategy used here was to employ several measures in order to trade their strengths and weaknesses off against one another and result in a comprehensive assessment of function. However, this is a cumbersome approach. Work needs to focus on identifying measures that better suit the needs of this patient group. Future studies could incorporate both quality of life and functioning so as to develop an understanding not only of the aspects that restrict someone's economic productivity, but also the factors that act as a barrier to greater life-satisfaction.

Although psychological factors were identified as significant predictors of psychosocial function, the mechanism by which they impact on functioning was not

investigated. Prospective longitudinal studies of self-esteem interventions in this group would help to identify 1) whether self-esteem interventions can be effective in this patient group, and 2) whether there is a causal relationship between increased self-esteem and improved functioning.

One further area which deserves exploration is the relationship between an individual's beliefs and fears about their cognitive function and their test performance. There are many factors that interact in a testing situation and impact on performance. Patients' concerns and subjective beliefs about their cognition and how it has changed over time is one. It was more common for patients to mention concerns and subjective beliefs about their cognition and how it has deteriorated over time. It is possible that factors such as this affect different patients to different degrees. The overall net effect is group differences in cognitive function, however it is not necessary or likely that each individual shows cognitive deficits for the same reasons. Exploring each of these aspects, and potentially many others, would provide an indicator of how much deficit remains to be explained.

REFERENCES

AIGNER, M., SACHS, G., BRUCKMULLER, E., WINKLBAUR, B., ZITTERL, W., KRYSPIN-EXNER, I., GUR, R. & KATSCHNIG, H. 2007. Cognitive and emotion recognition deficits in obsessive-compulsive disorder. *Psychiatry Research,* 149**,** 121-8.

AKISKAL, H. S., BOURGEOIS, M. L., ANGST, J., POST, R., MOLLER, H. & HIRSCHFELD, R. 2000. Re-evaluating the prevalence of and diagnostic composition within the broad clinical spectrum of bipolar disorders. *Journal of Affective Disorders,* 59**,** S5-S30.

ALLOY, L. B., ABRAMSON, L. Y., NEEREN, A. M., WALSHAW, P. D., UROSEVIC, S. & NUSSLOCK, R. 2006. Psychosocial risk factors for bipolar disorder: current and early enviornment and cognitive styles. *In:* JONES, S. & BENTALL, R. (eds.) *The Psychology of Bipolar Disorder: new developments and research strategies.* Oxford: Oxford University Press.

ALTMAN, E. G., HEDEKER, D., PETERSON, J. L. & DAVIS, J. M. 1997. The Altman Self-Rated Mania Scale. *Biological Psychiatry,* 42**,** 948-55.

ALTSHULER, L., MINTZ, J. & LEIGHT, K. 2002. The Life Functioning Questionnaire (LFQ): a brief, gender-neutral scale assessing functional outcome. *Psychiatry Research,* 112**,** 161-82.

ALTSHULER, L. L., BEARDEN, C. E., GREEN, M. F., VAN GORP, W. & MINTZ, J. 2008. A relationship between neurocognitive impairment and functional impairment in bipolar disorder: a pilot study. *Psychiatry Research,* 157**,** 289-93.

ALTSHULER, L. L., CURRAN, J. G., HAUSER, P., MINTZ, J., DENICOFF, K. & POST, R. 1995. T2 hyperintensities in bipolar disorder: magnetic resonance imaging comparison and literature meta-analysis. *American Journal of Psychiatry,* 152**,** 1139-44.

ANDERSON, M. C. & BJORK, R. A. 1994. Mechanisms of Inhibition in Long Term Memory: A New Taxonomy. *In:* DAGENBACH, D. & CARR, T. (eds.) *Inhibitory Processes in Attention, Memory & Language.* New York: Academic Press.

ANDREWS, G., PETERS, L. & TEESON, M. 1994. The Measurement of Consumer Outcome in Mental Health: A Report to the National Mental Health Information Strategy Committee. Sydney: Clinical Research Centre for Anxiety Disorders.

ANGST, J., GAMMA, A., BENAZZI, F., AJDACIC, V., EICH, D. & ROSSLER, W. 2003. Toward a re-definition of subthreshold bipolarity: epidemiology and proposed criteria for bipolar-II, minor bipolar disorders and hypomania. *Journal of Affective Disorders,* 73**,** 133-146.

ANTILA, M., TUULIO-HENRIKSSON, A., KIESEPPA, T., SORONEN, P., PALO, O. M., PAUNIO, T., HAUKKA, J., PARTONEN, T. & LONNQVIST, J. 2007. Heritability of cognitive

functions in families with bipolar disorder. *Am J Med Genet B Neuropsychiatr Genet,* 144B**,** 802-8.

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

ARNOLD, L. M., MCELROY, S. L. & KECK, P. E., JR. 2000. The role of gender in mixed mania. *Comprehensive Psychiatry,* 41**,** 83-7.

ARTS, B., JABBEN, N., KRABBENDAM, L. & VAN OS, J. 2008. Meta-analyses of cognitive functioning in euthymic bipolar patients and their first-degree relatives. *Psychological Medicine,* 38**,** 771-85.

ATRE-VAIDYA, N., TAYLOR, M. A., SEIDENBERG, M., REED, R., PERRINE, A. & GLICK-OBERWISE, F. 1998. Cognitive deficits, psychopathology, and psychosocial functioning in bipolar mood disorder. *Neuropsychiatry, Neuropsychology and Behavioral Neurology,* 11**,** 120-126.

ATTNEAVE, R. & ARNOULT, M. D. 1956. Methodological considerations in the quantitative study of shape and pattern in perception. *Psychological Bulletin,* 53**,** 452-471.

AUSTIN, M. P., MITCHELL, P., WILHELM, K., PARKER, G., HICKIE, I., BRODATY, H., CHAN, J., EYERS, K., MILIC, M. & HADZI-PAVLOVIC, D. 1999. Cognitive function in depression: a distinct pattern of frontal impairment in melancholia? *Psychol Med,* 29**,** 73-85.

BADDELEY, A. 1996. Exploring the central executive. *Quarterly Journal of Experimental Psychology,* 49A**,** 5-28.

BALANZA-MARTINEZ, V., TABARES-SEISDEDOS, R., SELVA-VERA, G., MARTINEZ-ARAN, A., TORRENT, C., SALAZAR-FRAILE, J., LEAL-CERCOS, C., VIETA, E. & GOMEZ-BENEYTO, M. 2005. Persistent cognitive dysfunctions in bipolar I disorder and schizophrenic patients: a 3-year follow-up study. *Psychotherapy & Psychosomatics,* 74**,** 113-9.

BALDASSANO, C. F., MARANGELL, L. B., GYULAI, L., NASSIR GHAEMI, S., JOFFE, H., KIM, D. R., SAGDUYU, K., TRUMAN, C. J., WISNIEWSKI, S. R., SACHS, G. S. & COHEN, L. S. 2005. Gender differences in bipolar disorder: retrospective data from the first 500 STEP-BD participants. *Bipolar Disorders,* 7**,** 465-70.

BALDESSARINI, R. J., TONDO, L. & HENNEN, J. 2003. Lithium treatment and suicide risk in major affective disorders: update and new findings. *Journal of Clinical Psychiatry,* 64**,** 44-52.

BARCH, D. M. 2009. Neuropsychological abnormalities in schizophrenia and major mood disorders: similarities and differences. *Current Psychiatry Reports,* 11**,** 313-9.

BARON-COHEN, S., WHEELWRIGHT, S., HILL, J., RASTE, Y. & PLUMB, I. 2001. The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry,* 42**,** 241-51.

BAUER, M. S., KIRK, G. F., GAVIN, C. & WILLIFORD, W. O. 2001. Determinants of functional outcome and healthcare costs in bipolar disorder: A high-intensity follow-up study. *Journal of Affective Disorders,* 65**,** 231-241.

BAUER, M. S., SIMON, G. E., LUDMAN, E. & UNUTZER, J. 2005. 'Bipolarity' in bipolar disorder: distribution of manic and depressive symptoms in a treated population. *British Journal of Psychiatry,* 187**,** 87-88.

BEARDEN, C. E., HOFFMAN, K. M. & CANNON, T. D. 2001. The neuropsychology and neuroanatomy of bipolar affective disorder: A critical review. *Bipolar Disorders,* 3**,** 106-150.

BECK, A. T. 1979. *Cognitive Therapy of Depression,* New York, Guilford Press.

BECK, A. T., WARD, C. H., MENDELSON, M., MOCK, J. & ERBAUGH, J. 1961. An inventory for measuring depression. *Archives of General Psychiatry,* 4**,** 561-571.

BEEBE, D. W., RIS, M. D. & DIETRICH, K. N. 2000. The relationship between CVLT-C process scores and measures of executive functioning: lack of support among community-dwelling adolescents. *Journal of Clinical and Experimental Neuropsychology,* 22**,** 779-92.

BELLACK, A. S., MUESER, K. T., MORRISON, R. L., TIERNEY, A. & PODELL, K. 1990. Remediation of cognitive deficits in schizophrenia. *American Journal of Psychiatry,* 147**,** 1650-5.

BENTALL, R. 2004. *Madness Explained: Psychosis and Human Nature,* London, Penguin.

BENTALL, R. P., KINDERMAN, P. & MANSON, K. 2005. Self-discrepancies in bipolar disorder: comparison of manic, depressed, remitted and normal participants. *Br J Clin Psychol,* 44**,** 457-73.

BENTALL, R. P. & THOMPSON, M. 1990. Emotional Stroop performance and the manic defence. *British Journal of Clinical Psychology,* 29**,** 235-7.

BORA, E., VAHIP, S., GONUL, A. S., AKDENIZ, F., ALKAN, M., OGUT, M. & ERYAVUZ, A. 2005. Evidence for theory of mind deficits in euthymic patients with bipolar disorder. *Acta Psychiatrica Scandinavica,* 112**,** 110-116.

BORA, E., YUCEL, M. & PANTELIS, C. 2009. Cognitive endophenotypes of bipolar disorder: a meta-analysis of neuropsychological deficits in euthymic patients and their first-degree relatives. *Journal of Affective Disorders,* 113**,** 1-20.

BOUSFIELD, A. K. & BOUSFIELD, W. A. 1966. Measurement of clustering and of sequential constancies in repeated free recall. *Psychological Reports,* 19**,** 935-942.

BOWIE, C. R., LEUNG, W. W., REICHENBERG, A., MCCLURE, M. M., PATTERSON, T. L., HEATON, R. K. & HARVEY, P. D. 2008. Predicting schizophrenia patients' real-world behavior with specific neuropsychological and functional capacity measures. *Biological Psychiatry,* 63**,** 505-11.

BOWIE, C. R., REICHENBERG, A., PATTERSON, T. L., HEATON, R. K. & HARVEY, P. D. 2006. Determinants of real-world functional performance in schizophrenia subjects: correlations with cognition, functional capacity, and symptoms. *American Journal of Psychiatry,* 163**,** 418-25.

BOZIKAS, V. P., TONIA, T., FOKAS, K., KARAVATOS, A. & KOSMIDIS, M. H. 2006. Impaired emotion processing in remitted patients with bipolar disorder. *Journal of Affective Disorders,* 91**,** 53-56.

BURDICK, K. E., BRAGA, R. J., GOLDBERG, J. F. & MALHOTRA, A. K. 2007. Cognitive dysfunction in bipolar disorder: future place of pharmacotherapy. *CNS Drugs,* 21**,** 971-81.

BURDICK, K. E., ENDICK, C. J. & GOLDBERG, J. F. 2005. Assessing cognitive deficits in bipolar disorder: are self-reports valid? *Psychiatry Research,* 136**,** 43-50.

BURGESS, P. W. & SHALLICE, T. 1997. *Hayling Sentence Completion Test,* Suffolk, England, Thames Valley Test Co. Ltd.

CARLSON, G. A., BROMET, E. J., DRIESSENS, C., MOJTABAI, R. & SCHWARTZ, J. E. 2002. Age at onset, childhood psychopathology, and 2-year outcome in psychotic bipolar disorder. *American Journal of Psychiatry,* 159**,** 307-309.

CASSIDY, F. & CARROLL, B. J. 2001. Frequencies of signs and symptoms in mixed and pure episodes of mania: implications for the study of manic episodes. *Progress in Neuropsychopharmacology and Biological Psychiatry,* 25**,** 659-65.

CASSIDY, F., FOREST, K., MURRY, E. & CARROLL, B. J. 1998. A factor analysis of the signs and symptoms of mania. *Archives of General Psychiatry,* 55**,** 27-32.

CEDRUS Superlab 2.0.4. California: Cedrus Corporation.

CEDRUS Superlab Pro 4.0. California: Cedrus Corporation.

CHEN, C.-H., LENNOX, B., JACOB, R., CALDER, A., LUPSON, V., BISBROWN-CHIPPENDALE, R., SUCKLING, J. & BULLMORE, E. 2006. Explicit and Implicit Facial Affect Recognition in Manic and Depressed States of Bipolar Disorder: A Functional Magnetic Resonance Imaging Study. *Biological Psychiatry,* 59**,** 31-39.

CHENGAPPA, K. N. R., HENNEN, J., BALDESSARINI, R. J., KUPFER, D. J., YATHAM, L. N., GERSHON, S., W BAKER, R. & TOHEN, M. 2005. Recovery and functional outcomes following olanzapine treatment for bipolar I mania. *Bipolar Disorders,* 7**,** 68-76.

CHRISTENSEN, H., GRIFFITHS, K., MACKINNON, A. & JACOMB, P. 1997. A quantitative review of cognitive deficits in depression and Alzheimer-type dementia. *J Int Neuropsychol Soc,* 3**,** 631-51.

CLARK, L., IVERSEN, S. D. & GOODWIN, G. M. 2001. A neuropsychological investigation of prefrontal cortex involvement in acute mania. *American Journal of Psychiatry,* 158**,** 1605-1611.

CLARK, L., IVERSEN, S. D. & GOODWIN, G. M. 2002. Sustained attention deficit in bipolar disorder. *British Journal of Psychiatry,* 180**,** 313-319.

CONUS, P., COTTON, S., ABDEL-BAKI, A., LAMBERT, M., BERK, M. & MCGORRY, P. D. 2006. Symptomatic and functional outcome 12 months after a first episode of psychotic mania: barriers to recovery in a catchment area sample. *Bipolar Disorders,* 8**,** 221-31.

CONUS, P. & MCGORRY, P. D. 2002. First-episode mania: A neglected priority for early intervention. *Australian & New Zealand Journal of Psychiatry,* 36**,** 158-172.

CORPORATION, T. P. 1997. Wechsler Adult Intelligence Scale III. San Antonio, TX: The Psychological Corporation.

CORPORATION, T. P. 2002. *WAIS-III/WMS-III Technical Manual,* San Antonio, TX, The Psychological Corporation.

CORYELL, W., SCHEFTNER, W., KELLER, M., ENDICOTT, J., MASER, J. & KLERMAN, G. L. 1993. The enduring psychosocial consequences of mania and depression. *American Journal of Psychiatry,* 150**,** 720-7.

CRADDOCK, N. & JONES, I. 1999. Genetics of bipolar disorder. *Journal of Medical Genetics,* 36**,** 585-94.

CRADDOCK, N. & JONES, I. 2001. Molecular genetics of bipolar disorder. *British Journal of Psychiatry Supplement,* 41**,** s128-33.

CRADDOCK, N. & SKLAR, P. 2009. Genetics of bipolar disorder: successful start to a long journey. *Trends Genet*.

DARWIN, C. 1872. *The Expression of the Emotions in Man and Animals,* London, John Murray.

DE JONG, A., GIEL, R., SLOOFF, C. J. & WIERSMA, D. 1985. Social disability and outcome in schizophrenic patients. *British Journal of Psychiatry,* 147**,** 631-6.

DECKERSBACH, T., MCMURRICH, S., OGUTHA, J., SAVAGE, C. R., SACHS, G. & RAUCH, S. L. 2004a. Characteristics of non-verbal memory impairment in bipolar disorder: the role of encoding strategies. *Psychological Medicine,* 34**,** 823-32.

DECKERSBACH, T., NIERENBERG, A. A., KESSLER, R., LUND, H. G., AMETRANO, R. M., SACHS, G., RAUCH, S. L. & DOUGHERTY, D. 2009. Cognitive Rehabilitation for Bipolar Disorder: An Open Trial for Employed Patients with Residual Depressive Symptoms. *CNS Neurosci Ther*.

DECKERSBACH, T., SAVAGE, C. R., REILLY-HARRINGTON, N., CLARK, L., SACHS, G. & RAUCH, S. L. 2004b. Episodic memory impairment in bipolar disorder and obsessive-compulsive disorder: The role of memory strategies. *Bipolar Disorders,* 6**,** 233-244.

DEPP, C. A., SAVLA, G. N., MOORE, D. J., PALMER, B. W., STRICKER, J. L., LEBOWITZ, B. D. & JESTE, D. V. 2008. Short-term course of neuropsychological abilities in middle-aged and older adults with bipolar disorder. *Bipolar Disorders,* 10**,** 684-90.

DICKERSON, F. B., BORONOW, J. J., STALLINGS, C. R., ORIGONI, A. E., COLE, S. & YOLKEN, R. H. 2004. Association between Cognitive Functioning and Employment Status of Persons with Bipolar Disorder. *Psychiatric Services,* 55**,** 54-58.

DINN, W. M., HARRIS, C. L., AYCICEGI, A., GREENE, P. B., KIRKLEY, S. M. & REILLY, C. 2004. Neurocognitive function in borderline personality disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry,* 28**,** 329-341.

DION, G. L., TOHEN, M., ANTHONY, W. A. & WATERNAUX, C. S. 1988. Symptoms and functioning of patients with bipolar disorder six months after hospitalization. *Hospital & Community Psychiatry,* 39**,** 652-657.

DITTMANN, S., SEEMULLER, F., SCHWARZ, M. J., KLEINDIENST, N., STAMPFER, R., ZACH, J., BORN, C., BERNHARD, B., FAST, K., GRUNZE, H., ENGEL, R. R. & SEVERUS, E. 2007. Association of cognitive deficits with elevated homocysteine levels in euthymic bipolar patients and its impact on psychosocial functioning: preliminary results. *Bipolar Disorders,* 9**,** 63-70.

DIXON, T., KRAVARITI, E., FRITH, C., MURRAY, R. M. & MCGUIRE, P. K. 2004. Effect of symptoms on executive function in bipolar illness. *Psychological Medicine,* 34**,** 811-21.

DUFF, K., SCHOENBERG, M. R., SCOTT, J. G. & ADAMS, R. L. 2005. The relationship between executive functioning and verbal and visual learning and memory. *Archives of Clinical Neuropsychology,* 20**,** 111-122.

DUNN, E. C., WEWIORSKI, N. J. & ROGERS, E. S. 2008. The meaning and importance of employment to people in recovery from serious mental illness: results of a qualitative study. *Psychiatr Rehabil J,* 32**,** 59-62.

EKMAN, P. & FRIESEN, W. V. 1976. *Pictures of Facial Affect,* Palo Alto, CA, Consulting Psychologists Press.

EKMAN, P. & FRIESEN, W. V. 2003. *Unmasking the Face,* Cambridge, MA, Malor.

FAGIOLINI, A., KUPFER, D. J., MASALEHDAN, A., SCOTT, J. A., HOUCK, P. R. & FRANK, E. 2005. Functional impairment in the remission phase of bipolar disorder. *Bipolar Disorders,* 7**,** 281-285.

FARAVELLI, C., DEGL'INNOCENTI, B. G., AIAZZI, L., INCERPI, G. & PALLANTI, S. 1990. Epidemiology of mood disorders: a community survey in Florence. *Journal of Affective Disorders,* 20**,** 135-141.

FERRIER, I. N., CHOWDHURY, R., THOMPSON, J. M., WATSON, S. & YOUNG, A. H. 2004. Neurocognitive function in unaffected first-degree relatives of patients with bipolar disorder: a preliminary report. *Bipolar Disorders,* 6**,** 319-323.

FERRIER, I. N., STANTON, B. R., KELLY, T. P. & SCOTT, J. 1999. Neuropsychological function in euthymic patients with bipolar disorder. *British Journal of Psychiatry,* 175**,** 246-251.

FIELD, A. 2000. *Discovering Statistics Using SPSS for Windows,* London, SAGE.

FIORAVANTI, M., CARLONE, O., VITALE, B., CINTI, M. E. & CLARE, L. 2005. A meta-analysis of cognitive deficits in adults with a diagnosis of schizophrenia. *Neuropsychol Rev,* 15**,** 73-95.

FIRST, M. B., SPITZER, R. L., WILLIAMS, J. B. W. & GIBBON, M. 1995. *Structured Clinical Interview for DSM-IV (SCID-IV),* New York, Biometrics Research Department, New York State Psychiatric Institute.

FORTY, L., JONES, L., JONES, I., SMITH, D. J., CAESAR, S., FRASER, C., GORDON-SMITH, K., HYDE, S. & CRADDOCK, N. 2009. Polarity at illness onset in bipolar I disorder and clinical course of illness. *Bipolar Disorders,* 11**,** 82-8.

FOSSATI, P., COYETTE, F., ERGIS, A.-M. & ALLILAIRE, J.-F. 2002. Influence of age and executive functioning on verbal memory of inpatients with depression. *Journal of Affective Disorders,* 68**,** 261-271.

FOSSATI, P., HARVEY, P.-O., LE BASTARD, G., ERGIS, A.-M., JOUVENT, R. & ALLILAIRE, J.-F. 2004. Verbal memory performance of patients with a first depressive episode and patients with unipolar and bipolar recurrent depression. *Journal of Psychiatric Research,* 38**,** 137-144.

FRENCH, C. C., RICHARDS, A. & SCHOLFIELD, E. J. 1996. Hypomania, anxiety and the emotional Stroop. *British Journal of Clinical Psychology,* 35**,** 617-26.

FUJIKAWA, T., YAMAWAKI, S. & TOUHOUDA, Y. 1995. Silent cerebral infarctions in patients with late-onset mania. *Stroke,* 26**,** 946-9.

GARNO, J. L., GOLDBERG, J. F., RAMIREZ, P. M. & RITZLER, B. A. 2005. Impact of childhood abuse on the clinical course of bipolar disorder. *British Journal of Psychiatry,* 186**,** 121-5.

GATHERCOLE, S. E. 1999. Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences,* 3**,** 410-419.

GERSHBERG, F. B. & SHIMAMURA, A. P. 1995. Impaired use of organizational strategies in free recall following frontal lobe damage. *Neuropsychologia,* 33**,** 1305-1333.

GETZ, G. E., SHEAR, P. K. & STRAKOWSKI, S. M. 2003. Facial affect recognition deficits in bipolar disorder. *Journal of the International Neuropsychological Society,* 9**,** 623-32.

GILDENGERS, A. G., MULSANT, B. H., BEGLEY, A., MAZUMDAR, S., HYAMS, A. V., REYNOLDS III, C. F., KUPFER, D. J. & BUTTERS, M. A. 2009. The longitudinal course of cognition in older adults with bipolar disorder. *Bipolar Disorders,* 11**,** 744-52.

GITLIN, M. J., SWENDSEN, J., HELLER, T. L. & HAMMEN, C. 1995. Relapse and impairment in bipolar disorder. *American Journal of Psychiatry,* 152**,** 1635-1640.

GLAHN, D. C., BARRETT, J., BEARDEN, C. E., MINTZ, J., GREEN, M. F., SERAP MONKUL, E., NAJT, P., SOARES, J. C. & VELLIGAN, D. I. 2006. Dissociable mechanisms for memory impairment in bipolar disorder and schizophrenia. *Psychological Medicine,* 36**,** 1085-1095.

GLAHN, D. C., BEARDEN, C. E., BARGUIL, M., BARRETT, J., REICHENBERG, A., BOWDEN, C. L., SOARES, J. C. & VELLIGAN, D. I. 2007. The neurocognitive signature of psychotic bipolar disorder. *Biological Psychiatry,* 62**,** 910-6.

GLAHN, D. C., BEARDEN, C. E., NIENDAM, T. A. & ESCAMILLA, M. A. 2004. The feasibility of neuropsychological endophenotypes in the search for genes associated with bipolar affective disorder. *Bipolar Disorders,* 6**,** 171-182.

GLASS, G. V., PECKHAM, P. D. & SANDERS, J. R. 1972. Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analyses of Variance and Covariance. *Review of Educational Research,* 42**,** 237-288.

GOELEVEN, E., DE RAEDT, R., LEYMAN, L. & VERSCHUERE, B. 2008. The Karolinska Directed Emotional Faces: A validation study. *Cognition & Emotion,* 22**,** 1094 - 1118.

GOETZ, I., TOHEN, M., REED, C., LORENZO, M. & VIETA, E. 2007. Functional impairment in patients with mania: baseline results of the EMBLEM study. *Bipolar Disorders,* 9**,** 45-52.

GOLDBERG, J. F. & GARNO, J. L. 2005. Development of posttraumatic stress disorder in adult bipolar patients with histories of severe childhood abuse. *Journal of Psychiatric Research,* 39**,** 595-601.

GOLDBERG, T. E. & WEINBERGER, D. R. 1994. Schizophrenia, training paradigms, and the Wisconsin Card Sorting Test redux. *Schizophrenia Research,* 11**,** 291-296.

GOLDMAN, H. H., SKODOL, A. E. & LAVE, T. R. 1992a. Revising axis V for DSM-IV: a review of measures of social functioning. *American Journal of Psychiatry,* 149**,** 1148-56.

GOLDMAN, R. S., AXELROD, B. N. & TOMPKINS, L. M. 1992b. Effect of instructional cues on schizophrenic patients' performance on the Wisconsin Card Sorting Test. *American Journal of Psychiatry,* 149**,** 1718-22.

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**,** 731-9.

GRAY, J., VENN, H., MONTAGNE, B., MURRAY, L., BURT, M., FRIGERIO, E., PERRETT, D. & YOUNG, A. H. 2006. Bipolar patients show mood-congruent biases in sensitivity to facial expressions of emotion when exhibiting depressed symptoms, but not when exhibiting manic symptoms. *Cogn Neuropsychiatry,* 11**,** 505-20.

GREEN, M. F. 1996. What are the functional consequences of neurocognitive deficits in schizophrenia? *American Journal of Psychiatry,* 153**,** 321-30.

GREEN, M. F., GANZELL, S., SATZ, P. & VACLAV, J. F. 1990. Teaching the Wisconsin Card Sorting Test to schizophrenic patients. *Archives of General Psychiatry,* 47**,** 91-2.

GREEN, M. F., KERN, R. S., BRAFF, D. L. & MINTZ, J. 2000. Neurocognitive deficits and functional outcome in schizophrenia: Are we measuring the "right stuff"? *Schizophrenia Bulletin,* 26**,** 119-136.

GUR, R. C., ERWIN, R. J., GUR, R. E., ZWIL, A. S., HEIMBERG, C. & KRAEMER, H. C. 1992. Facial emotion discrimination: II. Behavioral findings in depression. *Psychiatry Research,* 42**,** 241-51.

HAMILTON, M. 1960. A rating scale for depression. *Journal of Neurology Neurosurgery and Psychiatry,* 23**,** 56-62.

HARMER, C. J., GRAYSON, L. & GOODWIN, G. M. 2002. Enhanced recognition of disgust in bipolar illness. *Biological Psychiatry,* 51**,** 298-304.

HARRE, R. M. (ed.) 1986. *The Social Construction of Emotion,* Oxford: Blackwell.

HARVEY, K. E., GALLETLY, C. A., FIELD, C. & PROEVE, M. 2009. The effects of verbalisation on cognitive performance in schizophrenia: A pilot study using tasks from the Delis Kaplan Executive Function System. 1-9.

HASLER, G., DREVETS, W. C., GOULD, T. D., GOTTESMAN, I. I. & MANJI, H. K. 2006. Toward Constructing an Endophenotype Strategy for Bipolar Disorders. *Biological Psychiatry,* 60**,** 93-105.

HAYS, R. D., SHERBOURNE, C. D. & MAZEL, R. M. 1993. The RAND 36-Item Health Survey 1.0. *Health Economics,* 2**,** 217-27.

HEATON, R. K., CHELUNE, G. J., TALLEY, J. L., KAY, G. G. & CURTISS, G. 1993. *Wisconsin Card Sorting Test Manual,* Lutz, FL, Psychological Assessment Resources Inc.

HEIM, C., NEWPORT, D. J., BONSALL, R., MILLER, A. H. & NEMEROFF, C. B. 2001. Altered pituitary-adrenal axis responses to provocative challenge tests in adult survivors of childhood abuse. *American Journal of Psychiatry,* 158**,** 575-81.

HEINRICHS, R. W. & ZAKZANIS, K. K. 1998. Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. *Neuropsychology,* 12**,** 426-45.

HELLMAN, S. G., KERN, R. S., NEILSON, L. M. & GREEN, M. F. 1998. Monetary reinforcement and Wisconsin Card Sorting performance in schizophrenia: Why Show Me the Money? *Schizophrenia Research,* 34**,** 67-75.

HEUBROK, D. 1999. Subjective Organization of Verbal Memory and Learning in Adolescents with Brain Damage. *Child Neuropsychology,* 5**,** 24-33.

HO, T. P., GRAY, J., RATCLIFFE, A. A., REES, S., ROCKEY, J. & WIGHT, R. G. 2005. Does cognitive function influence alaryngeal speech rehabilitation? *Head Neck*.

IVLEVA, E. I., MORRIS, D. W., MOATES, A. F., SUPPES, T., THAKER, G. K. & TAMMINGA, C. A. 2010. Genetics and intermediate phenotypes of the schizophrenia--bipolar disorder boundary. *Neuroscience & Biobehavioral Reviews,* 34**,** 897-921.

IZARD, C. E. 1993. Four systems for emotion activation: cognitive and noncognitive processes. *Psychological Review,* 100**,** 68-90.

JOHNSON, S. L. & FINGERHUT, R. 2006. Life events as predictors of relapse, depression, and mania in bipolar disorder. *In:* JONES, S. & BENTALL, R. (eds.) *The Psychology of Bipolar Disorder: new developments and research strategies.* Oxford: Oxford University Press.

JONES, S. H. & BENTALL, R. P. 2008. A review of potential cognitive and environmental risk markers in children of bipolar parents. *Clinical Psychology Review,* 28**,** 1083-95.

JONGEN, E. M. M., SMULDERS, F. T. Y., RANSON, S. M. G., ARTS, B. M. G. & KRABBENDAM, L. 2007. Attentional bias and general orienting processes in bipolar disorder. *Journal of Behavior Therapy and Experimental Psychiatry,* 38**,** 168-83.

JUDD, L. L., AKISKAL, H. S., SCHETTLER, P. J., CORYELL, W., ENDICOTT, J., MASER, J. D., SOLOMON, D. A., LEON, A. C. & KELLER, M. B. 2003. A prospective investigation of the natural history of the long-term weekly symptomatic status of bipolar II disorder. *Archives of General Psychiatry,* 60**,** 261-9.

JUDD, L. L., AKISKAL, H. S., SCHETTLER, P. J., ENDICOTT, J., MASER, J., SOLOMON, D. A., LEON, A. C., RICE, J. A. & KELLER, M. B. 2002. The long-term natural history of the weekly symptomatic status of bipolar I disorder. *Archives of General Psychiatry,* 59**,** 530-7.

KATSYRI, J. & SAMS, M. 2008. The effect of dynamics on identifying basic emotions from synthetic and natural faces. *International Journal of Human-Computer Studies,* 66**,** 233-242.

KAY, J. H., ALTSHULER, L. L., VENTURA, J. & MINTZ, J. 1999. Prevalence of axis II comorbidity in bipolar patients with and without alcohol use disorders. *Annals of Clinical Psychiatry,* 11**,** 187-95.

KAZEN, J. K. & OTANI, H. 1996. An SPSS program to compute subjective organization. *Behavior Research Methods Instruments & Computers,* 28**,** 476-478.

KECK, P. E., MCELROY, S. L., STRAKOWSKI, S. M., WEST, S. A., SAX, K. W., HAWKINS, J. M., BOURNE, M. L. & HAGGARD, P. 1998. 12-month outcome of patients with bipolar disorder following hospitalization for a manic or mixed episode. *American Journal of Psychiatry,* 155**,** 646-652.

KEE, K. S., GREEN, M. F., MINTZ, J. & BREKKE, J. S. 2003. Is Emotion Processing a Predictor of Functional Outcome in Schizophrenia? *Schizophrenia Bulletin,* 29**,** 487-497.

KEMPER, T. D. 1987. How Many Emotions Are There? Wedding the Social and the Autonomic Components. *The American Journal of Sociology,* 93**,** 263-289.

KERI, S., KELEMEN, O., BENEDEK, G. & JANKA, Z. 2001. Different trait markers for schizophrenia and bipolar disorder: a neurocognitive approach. *Psychological Medicine,* 31**,** 915-22.

KERR, N., DUNBAR, R. I. M. & BENTALL, R. P. 2003. Theory of mind deficits in bipolar affective disorder. *Journal of Affective Disorders,* 73**,** 253-259.

KERR, N., SCOTT, J. & PHILLIPS, M. L. 2005. Patterns of attentional deficits and emotional bias in bipolar and major depressive disorder. *British Journal of Clinical Psychology,* 44**,** 343-356.

KESSING, L. V., HANSEN, M. G., ANDERSEN, P. K. & ANGST, J. 2004. The predictive effect of episodes on the risk of recurrence in depressive and bipolar disorders - a life-long perspective. *Acta Psychiatrica Scandinavica,* 109**,** 339-44.

KESSLER, R. C., RUBINOW, D. R., HOLMES, C., ABELSON, J. M. & ZHAO, S. 1997. The epidemiology of DSM-III-R bipolar I disorder in a general population survey. *Psychological Medicine,* 27**,** 1079-89.

KHAN, A., GINSBERG, L. D., ASNIS, G. M., GOODWIN, F. K., DAVIS, K. H., KRISHNAN, A. A. & ADAMS, B. E. 2004. Effect of lamotrigine on cognitive complaints in patients with bipolar I disorder. *Journal of Clinical Psychiatry,* 65**,** 1483-90.

KIESEPPÄ, T., TUULIO-HENRIKSSON, A., HAUKKA, J., VAN ERP, T., GLAHN, D., CANNON, T. D., PARTONEN, T., KAPRIO, J. & LÖNNQVIST, J. 2005. Memory and verbal learning functions in twins with bipolar-I disorder, and the role of information-processing speed. *Psychological Medicine,* 35**,** 205.

KILOH, L. G. 1962. Pseudo-dementia. *Acta Psychiatrica Scandinavica,* 38**,** 336-51.

KIM, E. Y. & MIKLOWITZ, D. J. 2004. Expressed emotion as a predictor of outcome among bipolar patients undergoing family therapy. *Journal of Affective Disorders,* 82**,** 343-52.

KNOWLES, R., TAI, S., JONES, S. H., HIGHFIELD, J., MORRISS, R. & BENTALL, R. P. 2007. Stability of self-esteem in bipolar disorder: comparisons among remitted bipolar patients, remitted unipolar patients and healthy controls. *Bipolar Disorders,* 9**,** 490-5.

KOSTER, E. H., DE RAEDT, R., GOELEVEN, E., FRANCK, E. & CROMBEZ, G. 2005. Mood-Congruent Attentional Bias in Dysphoria: Maintained Attention to and Impaired Disengagement From Negative Information. *Emotion,* 5**,** 446-455.

KOSTER, E. H. W., LEYMAN, L., RAEDT, R. D. & CROMBEZ, G. 2006. Cueing of visual attention by emotional facial expressions: The influence of individual differences in anxiety and depression. *Personality and Individual Differences,* 41**,** 329-339.

KURTZ, M. M. & GERRATY, R. T. 2009. A Meta-Analytic Investigation of Neurocognitive Deficits in Bipolar Illness: Profile and Effects of Clinical State. *Neuropsychology,* 23**,** 551-562.

LAES, J., SPONHEIM, S. R. & MACDONALD, A. W. 2005. Relationships between cognitive functions and current social functioning in schizophrenia and bipolar affective disorder. *Schizophrenia Bulletin,* 31**,** 364-364.

LAES, J. R. & SPONHEIM, S. R. 2006. Does cognition predict community function only in schizophrenia?: a study of schizophrenia patients, bipolar affective disorder patients, and community control subjects. *Schizophr Res,* 84**,** 121-31.

LAM, D. H., JONES, S. H., HAYWARD, P. & BRIGHT, J. A. 1999. *Cognitive therapy for bipolar disorder,* Chichester, Wiley.

LAZARUS, R. S. 1984. On the Primacy of Cognition. *American Psychologist,* 38**,** 124-129.

LEAHY, R. L. 1999. Decision making and mania. *Journal of Cognitive Psychotherapy: An International Quarterly,* 13**,** 83-105.

LEDOUX, J. 1998. *The Emotional Brain,* New York, Phoenix.

LEFF, J. 2001. *The Unbalanced Mind,* New York, Columbia University Press.

LEMBKE, A. & KETTER, T. A. 2002. Impaired recognition of facial emotion in mania. *American Journal of Psychiatry,* 159**,** 302-4.

LENNOX, B. R., JACOB, R., CALDER, A. J., LUPSON, V. & BULLMORE, E. T. 2004. Behavioural and neurocognitive responses to sad facial affect are attenuated in patients with mania. *Psychological Medicine,* 34**,** 795-802.

LEON, A. C., SOLOMON, D. A., MUELLER, T. I., ENDICOTT, J., POSTERNAK, M., JUDD, L. L., SCHETTLER, P. J., AKISKAL, H. S. & KELLER, M. B. 2000. A brief assessment of psychosocial functioning of subjects with bipolar I disorder: the LIFE-RIFT. Longitudinal Interval Follow-up Evaluation-Range Impaired Functioning Tool. *Journal of Nervous and Mental Disease,* 188**,** 805-12.

LEON, A. C., SOLOMON, D. A., MUELLER, T. I., TURVEY, C. L., ENDICOTT, J. & KELLER, M. B. 1999. The Range of Impaired Functioning Tool (LIFE-RIFT): a brief measure of functional impairment. *Psychological Medicine,* 29**,** 869-78.

LEVERICH, G. S., MCELROY, S. L., SUPPES, T., KECK, P. E., JR., DENICOFF, K. D., NOLEN, W. A., ALTSHULER, L. L., RUSH, A. J., KUPKA, R., FRYE, M. A., AUTIO, K. A. & POST, R. M. 2002. Early physical and sexual abuse associated with an adverse course of bipolar illness. *Biological Psychiatry,* 51**,** 288-97.

LEX, C., MEYER, T. D., MARQUART, B. & THAU, K. 2008. No strong evidence for abnormal levels of dysfunctional attitudes, automatic thoughts, and emotional information-processing biases in remitted bipolar I affective disorder. *Psychol Psychother,* 81**,** 1-13.

LEYMAN, L., DE RAEDT, R. & KOSTER, E. H. W. 2009. Attentional biases for emotional facial stimuli in currently depressed patients with bipolar disorder. *International Journal of Clinical and Health Psychology,* 9**,** 393-410.

LEYMAN, L., DE RAEDT, R., SCHACHT, R. I. K. & KOSTER, E. H. W. 2007. Attentional biases for angry faces in unipolar depression. *Psychological Medicine,* 37**,** 393-402.

LEZAK, M. D., HOWIESON, D. B. & LORING, D. W. 2004. *Neuropsychological Assessment,* Oxford, Oxford University Press.

LLOYD, T., KENNEDY, N., FEARON, P., KIRKBRIDE, J., MALLETT, R., LEFF, J., HOLLOWAY, J., HARRISON, G., DAZZAN, P., MORGAN, K., MURRAY, R. M. & JONES, P. B. 2005. Incidence of bipolar affective disorder in three UK cities: results from the AESOP study. *British Journal of Psychiatry,* 186**,** 126-31.

LOUIS, W. U. I. S. 2006. *English Lexicon Project Website* [Online]. Available: http://elexicon.wustl.edu/query14/query14.asp [Accessed].

LYON, H. M., STARTUP, M. & BENTALL, R. P. 1999. Social cognition and the manic defense: Attributions, selective attention, and self-schema in bipolar affective disorder. *Journal of Abnormal Psychology,* 108**,** 273-282.

MACKIN, P. & YOUNG, A. H. 2004. Rapid cycling bipolar disorder: historical overview and focus on emerging treatments. *Bipolar Disorders,* 6**,** 523-9.

MACKIN, P. & YOUNG, A. H. 2005. Bipolar Disorders. *In:* WRIGHT, P., STERN, J. & PHELAN, M. (eds.) *Core Psychiatry.* 2nd ed.: Saunders.

MACQUEEN, G. M., YOUNG, L. T. & JOFFE, R. T. 2001. A review of psychosocial outcome in patients with bipolar disorder. *Acta Psychiatrica Scandinavica,* 103**,** 163-170.

MAJ, M., MAGLIANO, L., PIROZZI, R., MARASCO, C. & GUARNERI, M. 1994. Validity of rapid cycling as a course specifier for bipolar disorder. *American Journal of Psychiatry,* 151**,** 1015-9.

MAJ, M., PIROZZI, R., FORMICOLA, A. M. & TORTORELLA, A. 1999. Reliability and validity of four alternative definitions of rapid-cycling bipolar disorder. *American Journal of Psychiatry,* 156**,** 1421-4.

MAJ, M., PIROZZI, R., MAGLIANO, L. & BARTOLI, L. 2002. The prognostic significance of 'switching' in patients with bipolar disorder: a 10-year prospective follow-up study. *American Journal of Psychiatry,* 159**,** 1711-1717.

MALHI, G. S., IVANOVSKI, B., HADZI-PAVLOVIC, D., MITCHELL, P. B., VIETA, E. & SACHDEV, P. 2007. Neuropsychological deficits and functional impairment in bipolar depression, hypomania and euthymia. *Bipolar Disorders,* 9**,** 114-25.

MALHI, G. S., LAGOPOULOS, J., SACHDEV, P. S., IVANOVSKI, B. & SHNIER, R. 2005. An emotional Stroop functional MRI study of euthymic bipolar disorder. *Bipolar Disorders,* 7**,** 58-69.

MANTERE, O., SUOMINEN, K., LEPPAMAKI, S., VALTONEN, H., ARVILOMMI, P. & ISOMETSA, E. 2004. The clinical characteristics of DSM-IV bipolar I and II disorders: baseline findings from the Jorvi Bipolar Study (JoBS). *Bipolar Disorders,* 6**,** 395-405.

MARINO, S. E., MEADOR, K. J., LORING, D. W., OKUN, M. S., FERNANDEZ, H. H., FESSLER, A. J., KUSTRA, R. P., MILLER, J. M., RAY, P. G., ROY, A., SCHOENBERG, M. R., VAHLE, V. J. & WERZ, M. A. 2009. Subjective perception of cognition is related to mood and not performance. *Epilepsy Behav*.

MARTINEZ-ARAN, A., TORRENT, C., TABARES-SEISDEDOS, R., SALAMERO, M., DABAN, C., BALANZA-MARTINEZ, V., SANCHEZ-MORENO, J., MANUEL GOIKOLEA, J., BENABARRE, A., COLOM, F. & VIETA, E. 2008. Neurocognitive impairment in bipolar patients with and without history of psychosis. *Journal of Clinical Psychiatry,* 69**,** 233-9.

MARTINEZ-ARAN, A., VIETA, E., COLOM, F., TORRENT, C., REINARES, M., GOIKOLEA, J. M., BENABARRE, A., COMES, M. & SANCHEZ-MORENO, J. 2005. Do cognitive complaints in euthymic bipolar patients reflect objective cognitive impairment? *Psychotherapy & Psychosomatics,* 74**,** 295-302.

MARTINEZ-ARAN, A., VIETA, E., COLOM, F., TORRENT, C., SANCHEZ-MORENO, J., REINARES, M., BENABARRE, A., GOIKOLEA, J. M., BRUGUE, E., DABAN, C. & SALAMERO, M. 2004a. Cognitive impairment in euthymic bipolar patients: Implications for clinical and functional outcome. *Bipolar Disorders,* 6**,** 224-232.

MARTINEZ-ARAN, A., VIETA, E., REINARES, M., COLOM, F., TORRENT, C., SANCHEZ-MORENO, J., BENABARRE, A., GOIKOLEA, J. M., COMES, M. & SALAMERO, M. 2004b. Cognitive Function Across Manic or Hypomanic, Depressed, and Euthymic States in Bipolar Disorder. *American Journal of Psychiatry,* 161**,** 262-270.

MARTINEZ-ARAN, A., VIETA, E., TORRENT, C., SANCHEZ-MORENO, J., GOIKOLEA, J. M., SALAMERO, M., MALHI, G. S., GONZALEZ-PINTO, A., DABAN, C., ALVAREZ-GRANDI, S., FOUNTOULAKIS, K., KAPRINIS, G., TABARES-SEISDEDOS, R. & AYUSO-MATEOS, J. L. 2007. Functional outcome in bipolar disorder: the role of clinical and cognitive factors. *Bipolar Disorders,* 9**,** 103-13.

MATHEWS, A. & MACLEOD, C. 1994. Cognitive Approaches to Emotion and Emotional Disorders. *Annual Review of Psychology,* 45**,** 25-50.

MCELROY, S. L., ALTSHULER, L. L., SUPPES, T., KECK, P. E., JR., FRYE, M. A., DENICOFF, K. D., NOLEN, W. A., KUPKA, R. W., LEVERICH, G. S., ROCHUSSEN, J. R., RUSH, A. J. & POST, R. M. 2001. Axis I psychiatric comorbidity and its relationship to historical illness variables in 288 patients with bipolar disorder. *American Journal of Psychiatry,* 158**,** 420-6.

MCGUFFIN, P., RIJSDIJK, F., ANDREW, M., SHAM, P., KATZ, R. & CARDNO, A. 2003. The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Archives of General Psychiatry,* 60**,** 497-502.

MCGURK, S. R., TWAMLEY, E. W., SITZER, D. I., MCHUGO, G. J. & MUESER, K. T. 2007. A meta-analysis of cognitive remediation in schizophrenia. *American Journal of Psychiatry,* 164**,** 1791-802.

MCHORNEY, C. A., WAR, J. E., JR., LU, J. F. R. & SHERBOURNE, C. D. 1994. The MOS 36-Item Short-Form Health Survey (SF-36): III. Tests of Data Quality, Scaling Assumptions, and Reliability across Diverse Patient Groups. *Medical Care,* 32**,** 40-66.

MCHORNEY, C. A., WARE, J. E., JR. & RACZEK, A. E. 1993. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and Clinical Tests of Validity in Measuring Physical and Mental Health Constructs. *Medical Care,* 31**,** 247-263.

MCKENNA, F. P. & SHARMA, D. 2004. Reversing the emotional Stroop effect reveals that it is not what it seems: the role of fast and slow components. *Journal of Experimental Psychology-Learning Memory and Cognition,* 30**,** 382-92.

MEYER, S. E., CARLSON, G. A., WIGGS, E. A., MARTINEZ, P. E., RONSAVILLE, D. S., KLIMES-DOUGAN, B., GOLD, P. W. & RADKE-YARROW, M. 2004. A prospective study of the association among impaired executive functioning, childhood attentional problems, and the development of bipolar disorder. *Developmental Psychopathology,* 16**,** 461-76.

MIKLOWITZ, D. J., WISNIEWSKI, S. R., MIYAHARA, S., OTTO, M. W. & SACHS, G. S. 2005. Perceived criticism from family members as a predictor of the one-year course of bipolar disorder. *Psychiatry Research,* 136**,** 101-11.

MILLER, G. A. 1956a. Human memory and the storage of information. *I.R.E. Transactions on Information Theory***,** 129-137.

MILLER, G. A. 1956b. The magical number seven, plus or minus two: Some limits on our capacity to process information. *Psychological Review,* 63**,** 81-97.

MITCHELL, P. B. & MALHI, G. S. 2004. Bipolar depression: phenomenological overview and clinical characteristics. *Bipolar Disorders,* 6**,** 530-539.

MIYAKE, A., FRIEDMAN, N. P., EMERSON, M. J., WITZKI, A. H., HOWERTER, A. & WAGER, T. D. 2000. The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology,* 41**,** 49-100.

MOGG, K. & BRADLEY, B. P. 2005. Attentional Bias in Generalized Anxiety Disorder Versus Depressive Disorder. *Cognitive Therapy and Research,* 29**,** 29-45.

MONARCH, E. S., SAYKIN, A. J. & FLASHMAN, L. A. 2004. Neuropsychological impairment in borderline personality disorder. *Psychiatric Clinics of North America,* 27**,** 67-+.

MONTAGNE, B., KESSELS, R. P., DE HAAN, E. H. & PERRETT, D. I. 2007. The Emotion Recognition Task: a paradigm to measure the perception of facial emotional expressions at different intensities. *Perceptual & Motor Skills,* 104**,** 589-98.

MOORE, P. B., SHEPHERD, D. J., ECCLESTON, D., MACMILLAN, I. C., GOSWAMI, U., MCALLISTER, V. L. & FERRIER, I. N. 2001. Cerebral white matter lesions in bipolar affective disorder: relationship to outcome. *British Journal of Psychiatry,* 178**,** 172-6.

MORRISS, R. 2002. Clinical importance of inter-episode symptoms in patients with bipolar affective disorder. *Journal of Affective Disorders,* 72**,** S3-13.

MULLER-OERLINGHAUSEN, B., BERGHOFER, A. & BAUER, M. 2002. Bipolar disorder. *Lancet,* 359**,** 241-7.

MUMMERY, C. J., PATTERSON, K., HODGES, J. R. & WISE, R. J. S. 1996. Generating 'tiger' as an animal name or a word beginning with T: Differences in brain activation. *Proceedings of the Royal Society of London Series B-Biological Sciences,* 263**,** 989-995.

MUR, M., PORTELLA, M. J., MARTINEZ-ARAN, A., PIFARRE, J. & VIETA, E. 2008. Long-term stability of cognitive impairment in bipolar disorder: a 2-year follow-up study of lithium-treated euthymic bipolar patients. *Journal of Clinical Psychiatry,* 69**,** 712-9.

MURPHY, F. C., SAHAKIAN, B. J., RUBINSZTEIN, J. S., MICHAEL, A., ROGERS, R. D., ROBBINS, T. W. & PAYKEL, E. S. 1999. Emotional bias and inhibitory control processes in mania and depression. *Psychological Medicine,* 29**,** 1307-1321.

MURPHY, M. D. & PUFF, C. R. 1982. Free recall: Basic methodology and analyses. *In:* PUFF, C. R. (ed.) *Handbook of Research Methods in Human Memory & Cognition.* New York: Academic Press.

NEHRA, R., CHAKRABARTI, S., PRADHAN, B. K. & KHEHRA, N. 2006. Comparison of cognitive functions between first- and multi-episode bipolar affective disorders. *Journal of Affective Disorders,* 93**,** 185-92.

NEITZ, J., SUMMERFELT, P. & NEITZ, M. *The Neitz Test of Color Vision Manual,* Los Angeles, CA, Western Psychological Services Publishers Distributors.

NELSON, H. E. 1982. *National Adult Reading Test, NART,* Windsor, Nelson Publishing Company.

NEWMAN, C. F., LEAHY, R. L., BECK, A. T., REILLY-HARRINGTON, N. & GYULAI, L. 2002. *Bipolar disorder: A cognitive therapy approach,* Washington, American Psychological Association.

ORGANISATION, W. H. 2001. *International Classification of Functioning, Disability and Health,* Geneva, World Health Organisation.

PARADISO, S., LAMBERTY, G. J., GARVEY, M. J. & ROBINSON, R. G. 1997. Cognitive impairment in the euthymic phase of chronic unipolar depression. *Journal of Nervous & Mental Disease,* 185**,** 748-754.

PATTERSON, T. L., GOLDMAN, S., MCKIBBIN, C. L., HUGHS, T. & JESTE, D. V. 2001a. UCSD Performance-Based Skills Assessment: development of a new measure of everyday functioning for severely mentally ill adults. *Schizophrenia Bulletin,* 27**,** 235-45.

PATTERSON, T. L., MOSCONA, S., MCKIBBIN, C. L., DAVIDSON, K. & JESTE, D. V. 2001b. Social skills performance assessment among older patients with schizophrenia. *Schizophrenia Research,* 48**,** 351-360.

PELLEGRINO, J. W. 1971. A general measure of organization in free recall for variable unit size and internal sequential consistency. *Behavior Research Methods & Instrumentation. Vol.,* 3**,** 241-246.

PELLEGRINO, J. W. & BATTIG, W. F. 1974. Relationships among higher order organizational measures and free recall. *Journal of Experimental Psychology,* 102**,** 463-472.

PELLEGRINO, J. W. & HUBERT, L. J. 1982. The analysis of organisation and structure in free recall. *In:* PUFF, C. R. (ed.) *Handbook of Research Methods in Human Memory & Cognition.* New York: Academic Press.

PERRY, W., POTTERAT, E. G. & BRAFF, D. L. 2001. Self-monitoring enhances Wisconsin Card Sorting Test performance in patients with schizophrenia: performance is improved by simply asking patients to verbalize their sorting strategy. *J Int Neuropsychol Soc,* 7**,** 344-52.

PERUGI, G., MICHELI, C., AKISKAL, H. S., MADARO, D., SOCCI, C., QUILICI, C. & MUSETTI, L. 2000. Polarity of the first episode, clinical characteristics, and course of manic depressive illness: a systematic retrospective investigation of 320 bipolar I patients. *Comprehensive Psychiatry,* 41**,** 13-8.

PETRIDES, M. & MILNER, B. 1982. Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia,* 20**,** 249-262.

PINI, S., DE QUEIROZ, V., PAGNIN, D., PEZAWAS, L., ANGST, J., CASSANO, G. B. & WITTCHEN, H.-U. 2005. Prevalence and burden of bipolar disorders in European countries. *European Neuropsychopharmacology,* 15**,** 425-434.

PITTAM, J. & SCHERER, K. R. 1993. Vocal Expression and Communication of Emotion. *In:* LEWIS, M. & HAVILAND, J. M. (eds.) *Handbook of Emotions.* New York: The Guilford Press.

POSNER, M. I., SNYDER, C. R. & DAVIDSON, B. J. 1980. Attention and the detection of signals. *Journal of Experimental Psychology,* 109**,** 160-74.

POST, R. M., DENICOFF, K. D., LEVERICH, G. S., ALTSHULER, L. L., FRYE, M. A., SUPPES, T. M., RUSH, A. J., KECK, P. E., JR., MCELROY, S. L., LUCKENBAUGH, D. A., POLLIO, C., KUPKA, R. & NOLEN, W. A. 2003. Morbidity in 258 bipolar outpatients followed for 1 year with daily prospective ratings on the NIMH life chart method. *Journal of Clinical Psychiatry,* 64**,** 680-90; quiz 738-9.

POST, R. M., LEVERICH, G. S., XING, G. & WEISS, R. B. 2001. Developmental vulnerabilities to the onset and course of bipolar disorder. *Development and Psychopathology,* 13**,** 581-98.

POWER, M. J. 2005. Psychological approaches to bipolar disorders: A theoretical critique. *Clinical Psychology Review,* 25**,** 1101-1122.

POWER, M. J., KATZ, R., MCGUFFIN, P., DUGGAN, C. F., LAM, D. & BECK, A. T. 1994. The Dysfunctional Attitude Scale (DAS): A Comparison of Forms A and B and Proposals for a New Subscaled Version. *Journal of Research in Personality,* 28**,** 263-276.

PUTMAN, P., SAEVARSSON, S. & VAN HONK, J. 2007. Hypomanic trait is associated with a hypovigilant automatic attentional response to social cues of danger. *Bipolar Disorders,* 9**,** 779-83.

RAMAKERS, I. H., VISSER, P. J., AALTEN, P., MAES, H. L., LANSDAAL, H. G., MEIJS, C. J., JOLLES, J. & VERHEY, F. R. 2010. The predictive value of memory strategies for Alzheimer's disease in subjects with mild cognitive impairment. *Archives of Clinical Neuropsychology,* 25**,** 71-7.

REGEER, E. J., TEN HAVE, M., ROSSO, M. L., HAKKAART-VAN ROIJEN, L., VOLLEBERGH, W. & NOLEN, W. A. 2004. Prevalence of bipolar disorder in the general population: a Reappraisal Study of the Netherlands Mental Health Survey and Incidence Study. *Acta Psychiatrica Scandinavica,* 110**,** 374-82.

REICHENBERG, A., WEISER, M., RABINOWITZ, J., CASPI, A., SCHMEIDLER, J., MARK, M., KAPLAN, Z. & DAVIDSON, M. 2002. A population-based cohort study of premorbid intellectual, language, and behavioral functioning in patients with schizophrenia, schizoaffective disorder, and nonpsychotic bipolar disorder. *American Journal of Psychiatry,* 159**,** 2027-2035.

REITAN, R. M. 1958. Validity of theTrail MakingTest as an indicator of organic brain damage. *Perceptual & Motor Skills,* 8**,** 271-276.

REY, A. 1964. *L'Examen Clinique en Psychologie,* Paris, Press Universitaire de France.

RIHMER, Z. & KISS, K. 2002. Bipolar disorders and suicidal behaviour. *Bipolar Disorders,* 4**,** 21-5.

ROBINSON, L. J. & FERRIER, I. N. 2006. Temporal Evolution of Cognitive Impairment in Bipolar Disorder: A Systematic Review of Cross Sectional Evidence. *Bipolar Disorders,* 8**,** 103-116.

ROBINSON, L. J., THOMPSON, J. M., GALLAGHER, P., GOSWAMI, U., YOUNG, A. H., FERRIER, I. N. & MOORE, P. B. 2006. A meta-analysis of cognitive deficits in euthymic patients with bipolar disorder. *Journal of Affective Disorders,* 93**,** 105-115.

ROSENBERG, E. & EKMAN, P. 1995. Conceptual and methodological issues in the judgment of facial expressions of emotion. *Motivation and Emotion,* 19**,** 111-138.

ROSENBERG, M. 1965. *Society and the Adolescent Self-Image,* Princeton, New Jersey, Princeton University Press.

ROSSER, A. & HODGES, J. R. 1994. Initial letter and semantic category fluency in Alzheimers Disease, Huntingtons Disease and Progressive Supranuclear palsy. *Journal of Neurology Neurosurgery and Psychiatry,* 57**,** 1389-1394.

RUBINSZTEIN, J. S., MICHAEL, A., UNDERWOOD, B. R., TEMPEST, M. & SAHAKIAN, B. J. 2006. Impaired cognition and decision-making in bipolar depression but no 'affective bias' evident. *Psychological Medicine,* 36**,** 629-639.

RUOCCO, A. C. 2005. The neuropsychology of borderline personality disorder: A meta-analysis and review. *Psychiatry Research,* 137**,** 191-202.

RUSSELL, J. A. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology,* 39**,** 1161-1178.

RUSSELL, J. A., BACHOROWSKI, J.-A. & FERNANDEZ-DOLS, J.-M. 2003. Facial and Vocal Expressions of Emotion. *Annual review of Psychology,* 54**,** 329-49.

RYBAKOWSKI, J. K. & TWARDOWSKA, K. 1999. The dexamethasone/corticotropin-releasing hormone test in depression in bipolar and unipolar affective illness. *Journal of Psychiatric Research,* 33**,** 363-370.

SAVITZ, J. B., SOLMS, M. & RAMESAR, R. S. 2005. Neurocognitive function as an endophenotype for genetic studies of bipolar affective disorder. *Neuromolecular Medicine,* 7**,** 275-86.

SCHERER, K. R. 2000. Psychological Models of Emotion. *In:* BOROD, J. C. (ed.) *The Neuropsychology of Emotion.* Oxford: Oxford University Press.

SCHERER, K. R. 2003. Vocal communication of emotion: A review of research paradigms. *Speech Communication,* 40**,** 227-256.

SCHMIDER, J., LAMMERS, C. H., GOTTHARDT, U., DETTLING, M., HOLSBOER, F. & HEUSER, I. J. 1995. Combined dexamethasone/corticotropin-releasing hormone test in acute and remitted manic patients, in acute depression, and in normal controls: I. *Biological Psychiatry,* 38**,** 797-802.

SCHNECK, C. D., MIKLOWITZ, D. J., CALABRESE, J. R., ALLEN, M. H., THOMAS, M. R., WISNIEWSKI, S. R., MIYAHARA, S., SHELTON, M. D., KETTER, T. A., GOLDBERG, J. F., BOWDEN, C. L. & SACHS, G. S. 2004. Phenomenology of rapid-cycling bipolar disorder: data from the first 500 participants in the Systematic Treatment Enhancement Program. *American Journal of Psychiatry,* 161**,** 1902-8.

SCHNEIDER, L. C. & STRUENING, E. L. 1983. SLOF: a behavioral rating scale for assessing the mentally ill. *Social Work Research and Abstracts,* 19**,** 9-21.

SEGALAS, C., ALONSO, P., LABAD, J., JAURRIETA, N., REAL, E., JIMENEZ, S., MENCHON, J. M. & VALLEJO, J. 2008. Verbal and nonverbal memory processing in patients with obsessive-compulsive disorder: its relationship to clinical variables. *Neuropsychology,* 22**,** 262-72.

SERRETTI, A., CAVALLINI, M. C., MACCIARDI, F., NAMIA, C., FRANCHINI, L., SOUERY, D., LIPP, O., BAUWENS, F., SMERALDI, E. & MENDLEWICZ, J. 1999. Social adjustment and self-esteem in remitted patients with mood disorders. *European Psychiatry,* 14**,** 137-142.

SHEEHAN, D. V., LECRUBIER, Y., SHEEHAN, K. H., AMORIM, P., JANAVS, J., WEILLER, E., HERGUETA, T., BAKER, R. & DUNBAR, G. C. 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry,* 59**,** 22-33; 34-57.

SHULMAN, K. I. 1997. Disinhibition syndromes, secondary mania and bipolar disorder in old age. *Journal of Affective Disorders,* 46**,** 175-82.

SIEMER, M. 2009. Mood experience: Implications of a dispositional theory of moods. *Emotion Review,* 1**,** 256-263.

SIMONSEN, C., SUNDET, K., VASKINN, A., UELAND, T., ROMM, K. L., HELLVIN, T., MELLE, I., FRIIS, S. & ANDREASSEN, O. A. 2010. Psychosocial function in schizophrenia and bipolar disorder: Relationship to neurocognition and clinical symptoms. *Journal of the International Neuropsychological Society***,** 1-13.

SOBCZAK, S., RIEDEL, W. J., BOOIJ, I., AAN HET ROT, M., DEUTZ, N. E. P. & HONIG, A. 2002. Cognition following acute tryptophan depletion: Difference between first-degree relatives of bipolar disorder patients and matched healthy control volunteers. *Psychological Medicine,* 32**,** 503-515.

SPIELBERGER, C. D. & GORSUCH, R. L. 1970. *STAI manual for the State-trait anxiety inventory*, Consulting Psychologists Press.

STERNBERG, R. J. & TULVING, E. 1977. The Measurement of Subjective Organisation in Free Recall. *Psychological Bulletin,* 84**,** 539-556.

STRAKOWSKI, S. M., WILLIAMS, J. R., FLECK, D. E. & DELBELLO, M. P. 2000. Eight-month functional outcome from mania following a first psychiatric hospitalization. *Journal of Psychiatric Research,* 34**,** 193-200.

STRAUSS, E., SHERMAN, E. M. S. & SPREEN, O. 2006. *A Compendium of Neuropsychological Tests: Administrations, Norms, and Commentary,* Oxford, Oxford University Press.

STRAUSS, M. E. & ALLRED, L. J. 1987. Measurement of Differential Cognitive Deficits after Head Injury. *In:* LEVIN, H. S., GRAFMAN, J. & EISENBERG, H. M. (eds.) *Neurobehavioural Recovery from Head Injury.* Oxford: Oxford University Press.

STROBER, M. & CARLSON, G. 1982. Bipolar illness in adolescents with major depression: clinical, genetic, and psychopharmacologic predictors in a three- to four-year prospective follow-up investigation. *Archives of General Psychiatry,* 39**,** 549-55.

STUDENMUND, A. H. 2001. *Using Econometrics: A Practical Guide,* Boston, Addison Wesley Longman Inc.

STUSS, D. T., ALEXANDER, M. P., PALUMBO, C. L., BUCKLE, L., SAYER, L. & POGUE, J. 1994. Organizational Strategies of Patients with Unilateral or Bilateral Frontal Lobe Injury in Word List Learning Tasks. *Neuropsychology,* 8**,** 355-373.

SUMMERFELT, A. T., ALPHS, L. D., FUNDERBURK, F. R., STRAUSS, M. E. & WAGMAN, A. M. 1991. Impaired Wisconsin Card Sort performance in schizophrenia may reflect motivational deficits. *Archives of General Psychiatry,* 48**,** 282-3.

SZADOCZKY, E., PAPP, Z., VITRAI, J., RIHMER, Z. & FUREDI, J. 1998. The prevalence of major depressive and bipolar disorders in Hungary. Results from a national epidemiologic survey. *Journal of Affective Disorders,* 50**,** 153-62.

SZOKE, A., SCHURHOFF, F., GOLMARD, J. L., ALTER, C., ROY, I., MEARY, A., ETAIN, B., BELLIVIER, F. & LEBOYER, M. 2006. Familial resemblance for executive functions in families of schizophrenic and bipolar patients. *Psychiatry Res,* 144**,** 131-8.

TEN HAVE, M., VOLLEBERGH, W., BIJL, R. & NOLEN, W. A. 2002. Bipolar disorder in the general population in The Netherlands (prevalence, consequences and care utilisation): results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Journal of Affective Disorders,* 68**,** 203-13.

THOMPSON, J. M., GALLAGHER, P., HUGHES, J. H., WATSON, S., GRAY, J. M., FERRIER, I. N. & YOUNG, A. H. 2005. Neurocognitive impairment in euthymic patients with bipolar affective disorder. *British Journal of Psychiatry,* 186**,** 32-40.

THOMPSON, J. M., GRAY, J. M., HUGHES, J. H., WATSON, S., YOUNG, A. H. & FERRIER, I. N. 2007. Impaired working memory monitoring in euthymic bipolar patients. *Bipolar Disorders,* 9**,** 478-89.

THOMPSON, J. M., HAMILTON, C. J., GRAY, J. M., QUINN, J. G., MACKIN, P., YOUNG, A. H. & FERRIER, I. N. 2006. Executive and visuospatial sketchpad resources in euthymic bipolar disorder: Implications for visuospatial working memory architecture. *Memory,* 14**,** 437-51.

THOMPSON, K. N., CONUS, P. O., WARD, J. L., PHILLIPS, L. J., KOUTSOGIANNIS, J., LEICESTER, S. & MCGORRY, P. D. 2003. The initial prodrome to bipolar affective disorder: prospective case studies. *Journal of Affective Disorders,* 77**,** 79-85.

TIIHONEN, J., HAUKKA, J., HENRIKSSON, M., CANNON, M., KIESEPPA, T., LAAKSONEN, I., SINIVUO, J. & LONNQVIST, J. 2005. Premorbid Intellectual Functioning in Bipolar Disorder and Schizophrenia: Results From a Cohort Study of Male Conscripts. *American Journal of Psychiatry,* 162**,** 1904-1910.

TOHEN, M., STRAKOWSKI, S. M., ZARATE, C., JR., HENNEN, J., STOLL, A. L., SUPPES, T., FAEDDA, G. L., COHEN, B. M., GEBRE-MEDHIN, P. & BALDESSARINI, R. J. 2000. The McLean-Harvard First-Episode Project: 6-month symptomatic and functional outcome in affective and nonaffective psychosis. *Biological Psychiatry. Special A special issue on bipolar disorder,* 48**,** 467-476.

TORRES, I. J., BOUDREAU, V. G. & YATHAM, L. N. 2007. Neuropsychological functioning in euthymic bipolar disorder: a meta-analysis. *Acta Psychiatrica Scandinavica, Supplementum***,** 17-26.

TREMONT, G., HALPERT, S., JAVORSKY, D. J. & STERN, R. A. 2000. Differential impact of executive dysfunction on verbal list learning and story recall. *Clinical Neuropsychologist,* 14**,** 295-302.

TSAI, S.-Y., CHEN, C.-C. & YEH, E.-K. 1997. Alcohol problems and long-term psychosocial outcome in Chinese patients with bipolar disorder. *Journal of Affective Disorders,* 46**,** 143-150.

TSAI, S.-Y. M., CHEN, C.-C., KUO, C.-J., LEE, J.-C., LEE, H.-C. & STRAKOWSKI, S. M. 2001. 15-year outcome of treated bipolar disorder. *Journal of Affective Disorders,* 63**,** 215-220.

TSUANG, M. T., WOOLSON, R. F. & FLEMING, J. A. 1979. Long-term outcome of major psychoses. I. Schizophrenia and affective disorders compared with psychiatrically symptom-free surgical conditions. *Archives of General Psychiatry,* 36**,** 1295-301.

TULVING, E. 1962. Subjective organization in free recall of "unrelated" words. *Psychological Review,* 69**,** 344-354.

TULVING, E. 1964. Intratrial and intertrial retention: Notes towards a theory of free recall verbal learning. *Psychological Review,* 71**,** 219-237.

VAN DER GAAG, M., KERN, R. S., VAN DEN BOSCH, R. J. & LIBERMAN, R. P. 2002. A controlled trial of cognitive remediation in schizophrenia. *Schizophrenia Bulletin,* 28**,** 167-76.

VAN GORP, W. G., ALTSHULER, L., THEBERGE, D. C., WILKINS, J. & DIXON, W. 1998. Cognitive impairment in euthymic bipolar patients with and without prior alcohol dependence: A preliminary study. *Archives of General Psychiatry,* 55**,** 41-46.

VASKINN, A., SUNDET, K., FRIIS, S., SIMONSEN, C., BIRKENAES, A. B., ENGH, J. A., JONSDOTTIR, H., RINGEN, P. A., OPJORDSMOEN, S. & ANDREASSEN, O. A. 2007. The effect of gender on emotion perception in schizophrenia and bipolar disorder. *Acta Psychiatrica Scandinavica,* 116**,** 263-70.

VENN, H. R., GRAY, J. M., MONTAGNE, B., MURRAY, L. K., MICHAEL BURT, D., FRIGERIO, E., PERRETT, D. I. & YOUNG, A. H. 2004. Perception of facial expressions of emotion in bipolar disorder. *Bipolar Disorders,* 6**,** 286-93.

VYGOTSKY, L. S. 1962. *Thought and language,* Cambridge, MA, MIT.

WARE, J. E., JR. & SHERBOURNE, C. D. 1992. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care,* 30**,** 473-83.

WATSON, S., GALLAGHER, P., RITCHIE, J. C., FERRIER, I. N. & YOUNG, A. H. 2004. Hypothalamic-pituitary-adrenal axis function in patients with bipolar disorder. *British Journal of Psychiatry,* 184**,** 496-502.

WECHSLER, D. 1997. *Wechsler Memory Scale—Third Edition (WMS-III),* San Antonio, TX, Psychological Corporation.

WEISSMAN, M. M. 1999. Social Adjustment Scale - Self Report (SAS-SR) User's Manual. New York: Multi-Health Systems.

WEISSMAN, M. M., BLAND, R. C., CANINO, G. J., FARAVELLI, C., GREENWALD, S., HWU, H. G., JOYCE, P. R., KARAM, E. G., LEE, C. K., LELLOUCH, J., LEPINE, J. P., NEWMAN, S. C., RUBIO-STIPEC, M., WELLS, J. E., WICKRAMARATNE, P. J., WITTCHEN, H. & YEH, E. K. 1996. Cross-national epidemiology of major depression and bipolar disorder. *Journal of the American Medical Association,* 276**,** 293-9.

WEISSMAN, M. M. & MYERS, J. K. 1978. Affective disorders in a US urban community: the use of research diagnostic criteria in an epidemiological survey. *Archives of General Psychiatry,* 35**,** 1304-11.

WEISSMAN, M. M., PRUSOFF, B. A., THOMPSON, W. D., HARDING, P. S. & MYERS, J. K. 1978. Social adjustment by self-report in a community sample and in psychiatric outpatients. *Journal of Nervous and Mental Disease,* 166**,** 317-26.

WILSON, B. A., ALDERMAN, N., BURGESS, P. W., EMSLIE, H. & EVANS, J. J. 1996. *Behavioural Assessment of the Dysexecutive Syndrome,* London, Harcourt Assessment.

WINTERS, K. C. & NEALE, J. M. 1985. Mania and low self-esteem. *Journal of Abnormal Psychology,* 94**,** 282-290.

WYKES, T., NEWTON, E., LANDAU, S., RICE, C., THOMPSON, N. & FRANGOU, S. 2007. Cognitive remediation therapy (CRT) for young early onset patients with schizophrenia: an exploratory randomized controlled trial. *Schizophrenia Research,* 94**,** 221-30.

YAN, L. J., HAMMEN, C., COHEN, A. N., DALEY, S. E. & HENRY, R. M. 2004. Expressed emotion versus relationship quality variables in the prediction of recurrence in bipolar patients. *Journal of Affective Disorders,* 83**,** 199-206.

YANOS, P. T., PRIMAVERA, L. H. & KNIGHT, E. L. 2001. Consumer-run service participation, recovery of social functioning, and the mediating role of psychological factors. *Psychiatric Services,* 52**,** 493-500.

YOUNG, R., BIGGS, J., ZIEGLER, V. & MEYER, D. 1978. A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry,* 133**,** 429-435.

ZAJONC, R. B. 1980. Feeling and thinking: preferences need no inferences. *American Psychologist,* 35**,** 151-175.

ZAJONC, R. B. 1984. On the Primacy of Affect. *American Psychologist,* 39**,** 117-123.

ZAKZANIS, K. K., LEACH, L. & KAPLAN, E. 1998. On the nature and pattern of neurocognitive function in major depressive disorder. *Neuropsychiatry Neuropsychology and Behavioral Neurology,* 11**,** 111-119.

ZALLA, T., JOYCE, C., SZOKE, A., SCHURHOFF, F., PILLON, B., KOMANO, O., PEREZ-DIAZ, F., BELLIVIER, F., ALTER, C., DUBOIS, B., ROUILLON, F., HOUDE, O. & LEBOYER, M. 2004. Executive dysfunctions as potential markers of familial vulnerability to bipolar disorder and schizophrenia. *Psychiatry Research,* 121**,** 207-217.

ZAMMIT, S., ALLEBECK, P., DAVID, A. S., DALMAN, C., HEMMINGSSON, T., LUNDBERG, I. & LEWIS, G. 2004. A longitudinal study of premorbid IQ score and risk of developing schizophrenia, bipolar disorder, severe depression, and other nonaffective psychoses. *Archives of General Psychiatry,* 61**,** 354-360.

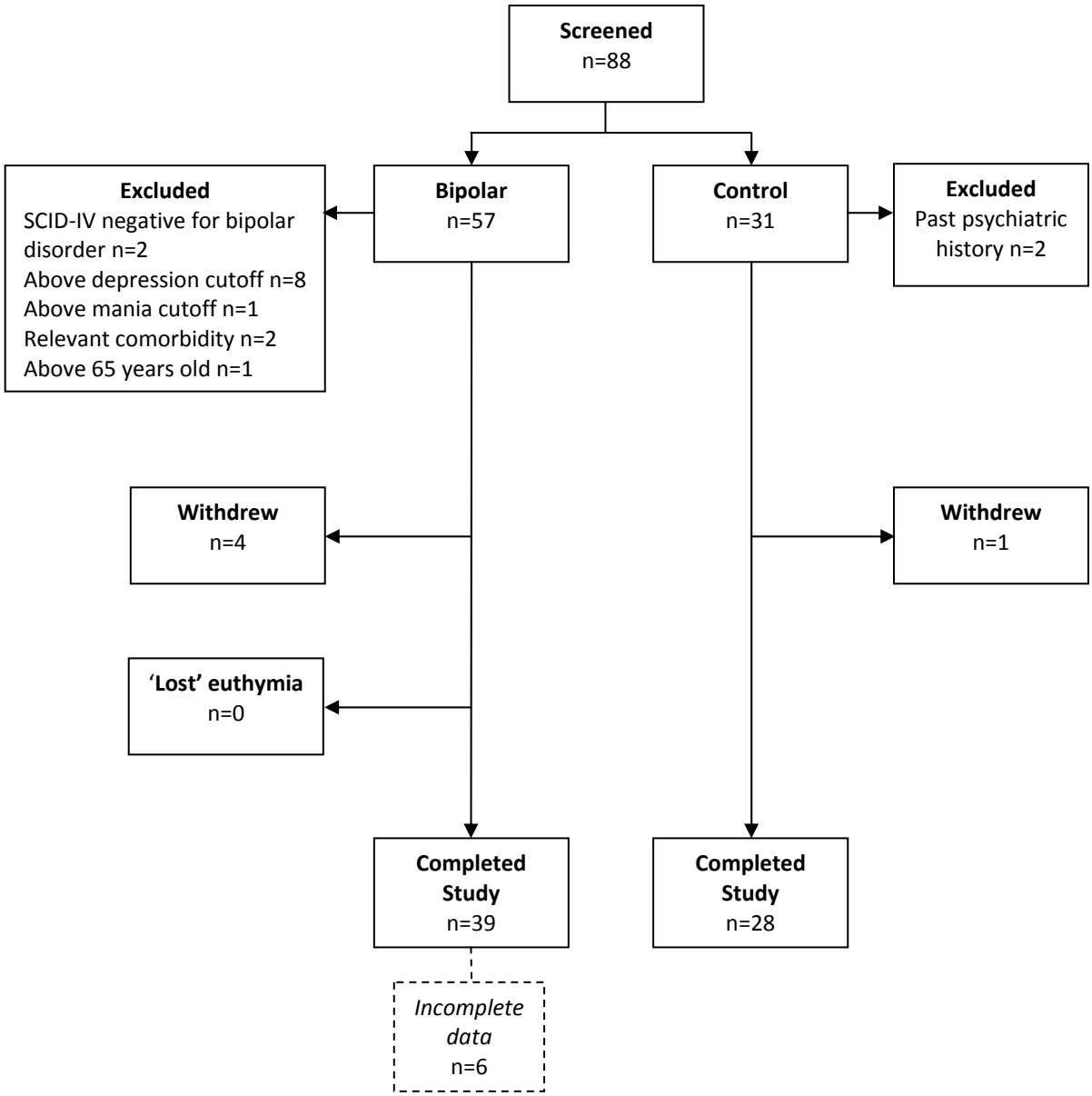
ZUBIETA, J. K., HUGUELET, P., O'NEIL, R. L. & GIORDANI, B. J. 2001. Cognitive function in euthymic Bipolar I Disorder. *Psychiatry Research,* 102**,** 9-20.

APPENDICES

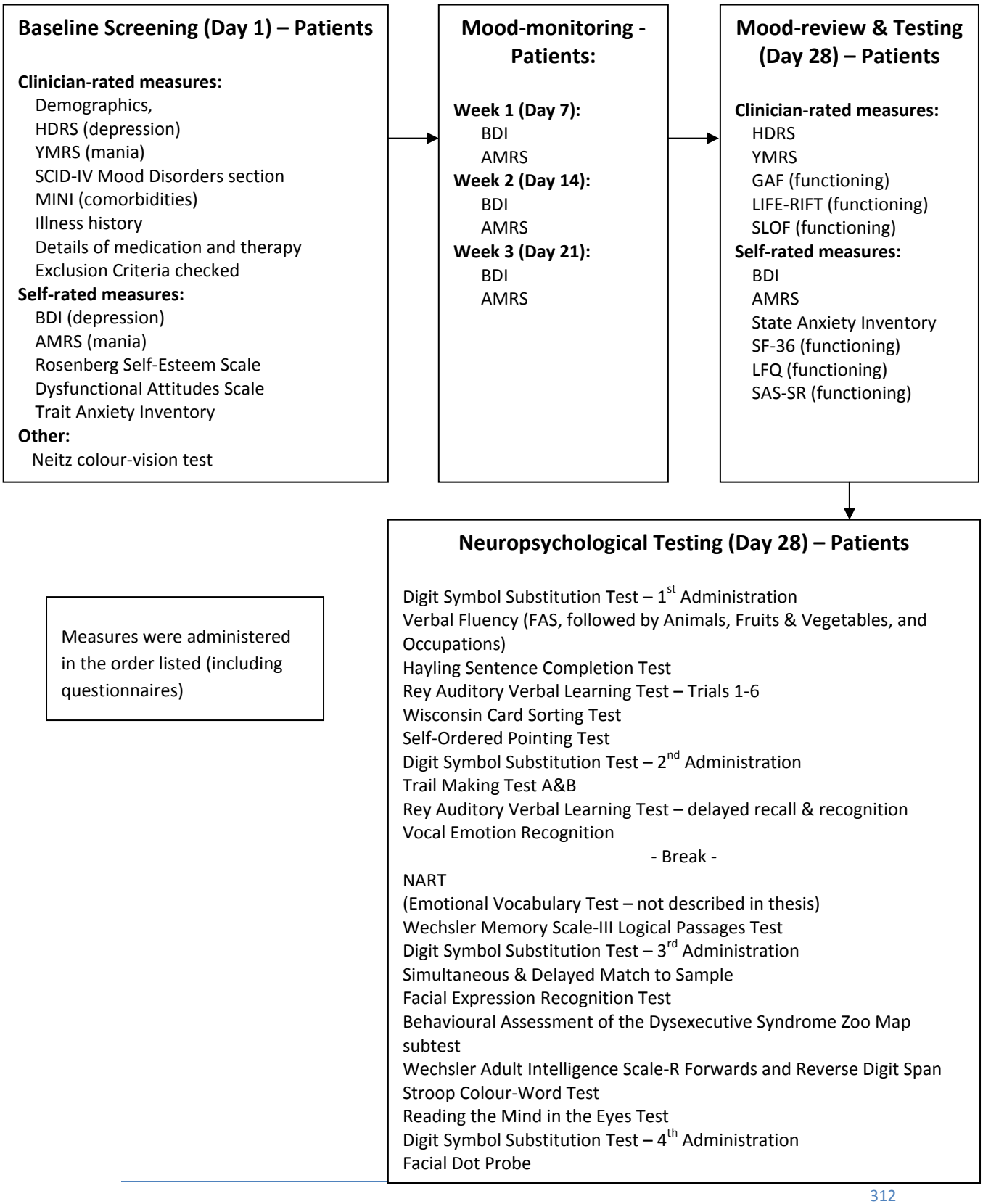
APPENDIX 1: LITERATURE SEARCH CRITERIA

To identify relevant studies of cognitive function in euthymic patients with bipolar disorder, the title and abstract fields of the electronic databases Medline (1966-present) and EMBASE (1980-present) were searched with the following terms: (Bipolar Disorder OR Manic Depress*) and (cognit* OR memory OR attention OR executive OR neurocognit* OR neuropsych*). (* indicates a wild card) Studies were deemed suitable if they met all of the following criteria: 1) included a group of adult patients diagnosed with bipolar disorder, 2) diagnosis was made by a recognised criterion-based diagnostic system (e.g. ICD-10, DSM-III, DSM-III-R, or DSM-IV), 3) included a healthy control group or normative comparison sample, 4) used recognised or experimental tests of neuropsychological function, but not merely measures of general cognition (e.g. Mini-Mental State Exam).

APPENDIX 2: FLOW CHART OF PARTICIPANTS



APPENDIX 3: STUDY PROCEDURE AND ORDER OF ADMINISTRATION



Baseline Screening (Day 1) – Controls

Clinician-rated measures:
Demographics,
HDRS (depression)
YMRS (mania)
SCID-IV Mood Disorders section
MINI
Exclusion Criteria checked
GAF (functioning)
LIFE-RIFT (functioning)
SLOF (functioning)

Self-rated measures:
BDI (depression)
AMRS (mania)
Rosenberg Self-Esteem Scale
Dysfunctional Attitudes Scale
Trait Anxiety Inventory

Other:
Neitz colour-vision test

Testing (Day 7) – Controls

Self-rated measures:
BDI
AMRS
State Anxiety Inventory
SF-36 (functioning)
LFQ (functioning)
SAS-SR (functioning)

Neuropsychological Testing (Day 7) – Controls

Digit Symbol Substitution Test – 1st Administration
Verbal Fluency (FAS, followed by Animals, Fruits & Vegetables, and Occupations)
Hayling Sentence Completion Test
Rey Auditory Verbal Learning Test – Trials 1-6
Wisconsin Card Sorting Test
Self-Ordered Pointing Test
Digit Symbol Substitution Test – 2nd Administration
Trail Making Test A&B
Rey Auditory Verbal Learning Test – delayed recall & recognition
Vocal Emotion Recognition
- Break -
NART
(Emotional Vocabulary Test – not described in thesis)
Wechsler Memory Scale-III Logical Passages Test
Digit Symbol Substitution Test – 3rd Administration
Simultaneous & Delayed Match to Sample
Facial Expression Recognition Test
Behavioural Assessment of the Dysexecutive Syndrome Zoo Map subtest
Wechsler Adult Intelligence Scale-R Forwards and Reverse Digit Span
Stroop Colour-Word Test
Reading the Mind in the Eyes Test
Digit Symbol Substitution Test – 4th Administration
Facial Dot Probe

Measures were administered in the order listed (including questionnaires)

APPENDIX 4: NEUROPSYCHOLOGICAL IMPAIRMENT COVARYING FOR IQ AND EDUCATION

Table A4.51: Results of ANCOVA including NART IQ as a covariate. The table includes the effect size from the ANOVA without the covariate, followed by the Pearson correlation between NART IQ and the neuropsychological test index score, followed by the significance level of the covariate from the ANCOVA, followed by the significance level of the group difference on the neuropsychological measure from the ANCOVA and finally the effect size of the group difference accounting for the covariate.

Covarying for IQ	Pre-covariate	Correlation			Covariate		Group		Post-covariate
	d	n	r	p	$F_{1,62}$	p	$F_{1,62}$	p	d
MEMORY - VERBAL									
Rey Auditory Verbal Learning Test									
Total recall trials 1-5	0.50	63	0.32	0.010	5.54[b]	0.02	2.52[b]	0.12	0.39
A1	0.56	64	0.39	0.002	8.98[c]	<0.01	2.93[c]	0.09	0.42
A6 (short delay)	0.42	64	0.26	0.036	3.53[c]	0.07	2.04[c]	0.16	0.35
Wechsler Memory Scale Logical Passages Test									
A+B Recall Units	0.72	63	0.30	0.019	3.85[b]	0.05	5.78[b]	0.02	0.60
A+B Thematic units	0.52	63	0.33	0.009	5.41[b]	0.02	2.93[b]	0.09	0.42
Forward Digit Span									
Span	0.21	64	0.29	0.021	5.10[c]	0.03	0.12[c]	0.73	0.09
Score	0.14	64	0.32	0.010	6.75[c]	0.01	0.00[c]	0.95	-0.02
EXECUTIVE - FLUENCY									
FAS									
Total correct	0.39	64	0.58	<0.001	28.83[c]	<0.01	0.51[c]	0.48	0.18
Category									
Total correct	0.72	65	0.40	0.001	9.06	<0.01	6.73	0.01	0.64
Total repeats	-0.16	65	-0.34	0.005	10.50	<0.01	3.28	0.08	0.44
EXECUTIVE - INHIBITION									
Hayling Sentence Completion Test									
Total time part 2 (scaled)	0.42	65	0.25	0.042	3.15	0.08	1.89	0.17	0.34
B Score	0.45	65	-0.30	0.014	4.87	0.03	2.27	0.14	0.37
Overall Scaled Score	0.31	65	0.32	0.009	6.18	0.02	0.69	0.41	0.20

Table A4:49 continued

Covarying for IQ	Pre-covariate	Correlation			Covariate		Group		Post-covariate
	d	n	r	p	$F_{1,62}$	p	$F_{1,62}$	p	d
EXECUTIVE - PLANNING									
BADS Zoo Map									
Version 1 total errors	0.51	60	-0.36	0.005	7.28[a]	0.01	3.23[a]	0.08	0.45
Self-Ordered Pointing Test									
Total errors all levels	0.62	65	-0.39	0.001	9.01	0.00	3.16	0.08	0.44
Total errors level 6	0.53	65	-0.32	0.009	5.64	0.02	2.23	0.14	0.37
Total errors level 8	0.50	65	-0.41	0.001	10.42	0.00	2.02	0.16	0.35
Total errors level 10	0.57	65	-0.31	0.011	5.20	0.03	2.60	0.11	0.40
Highest level with at least 1 trial correct	0.35	65	0.44	<0.001	13.33	<0.01	0.45	0.50	0.16
Maximum span level 6	0.29	65	0.36	0.003	8.27	0.01	0.40	0.53	0.16
Maximum span level 8	0.52	65	0.36	0.004	7.10	0.01	2.73	0.10	0.40
EXECUTIVE – MENTAL MANIPULATION									
Reverse Digit Span									
Span	0.32	64	0.46	<0.001	15.04[c]	<0.01	0.27[c]	0.60	0.13
Score	0.23	64	0.49	<0.001	18.69[c]	<0.01	0.01[c]	0.92	0.02
Total score (Fwd + Reverse)	0.20	64	0.46	<0.001	15.83[c]	<0.01	<0.01[c]	1.00	0.00
EXECUTIVE – SET-SHIFTING									
Trail Making Test									
TMTb time	0.20	64	-0.32	0.01	6.37[c]	0.01	0.05[c]	0.82	0.06
Switch (TMTa – TMTb)	0.01	64	-0.27	0.03	5.08[c]	0.03	0.20[c]	0.66	0.11
PSYCHOMOTOR									
Digit Symbol Substitution Test									
DSST 1	0.52	65	0.31	0.013	4.90	0.03	2.46	0.12	0.39
DSST 2	0.85	65	0.31	0.012	4.39	0.04	8.89	<0.01	0.73
DSST % change 2-3	-0.25	63	-0.41	0.001	11.10[b]	<0.01	0.21[b]	0.65	-0.11

[a] $F_{1,57}$; [b]$F_{1,60}$; [c]$F_{1,61}$

Table A4.52: Results of ANCOVA including years of education as a covariate. The table includes the effect size from the ANOVA without the covariate, followed by the Pearson correlation between years of education and the neuropsychological test index score, followed by the significance level of the covariate from the ANCOVA, followed by the significance level of the group difference on the neuropsychological measure from the ANCOVA and finally the effect size of the group difference accounting for the covariate.

Covarying for Education	Pre-covariate	Correlation			Covariate		Group		Post-covariate
	d	n	r	p	$F_{1,62}$	p	$F_{1,62}$	p	d
EXECUTIVE – FLUENCY									
FAS									
Total correct	0.39	65	0.39	0.001	9.51	<0.01	0.96	0.33	0.24
Category									
Total correct	0.72	66	0.29	0.019	3.78	0.06	6.45	0.01	0.62
EXECUTIVE - PLANNING									
BADS Zoo Map									
Version 1 total errors	0.51	61	-0.29	0.024	4.28[b]	0.04	2.69[b]	0.11	0.41
Overall profile score	0.44	60	0.28	0.032	4.03[a]	0.05	1.77[a]	0.19	0.34
Self-Ordered Pointing Test									
Total errors level 8	0.50	66	-0.27	0.029	3.65	0.06	2.17	0.15	0.36
Highest level with at least 1 trial correct	0.35	66	0.33	0.007	6.55	0.01	0.63	0.43	0.19
Maximum span level 6	0.29	66	0.30	0.016	5.23	0.03	0.41	0.53	0.16
EXECUTIVE – MENTAL MANIPULATION									
Reverse Digit Span									
Score	0.23	65	0.28	0.022	4.83	0.03	0.29	0.59	0.13

[a] $F_{1,57}$; [b] $F_{1,58}$

APPENDIX 5: RECALL & SUBJECTIVE ORGANISATION GRAPHS

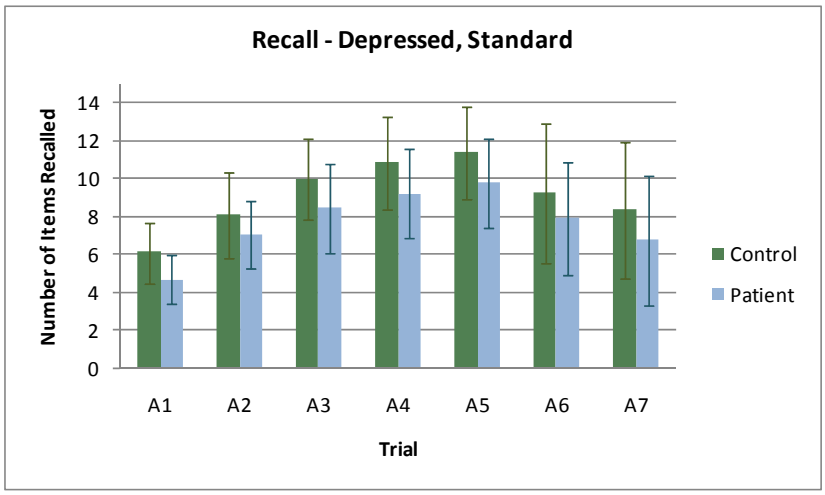
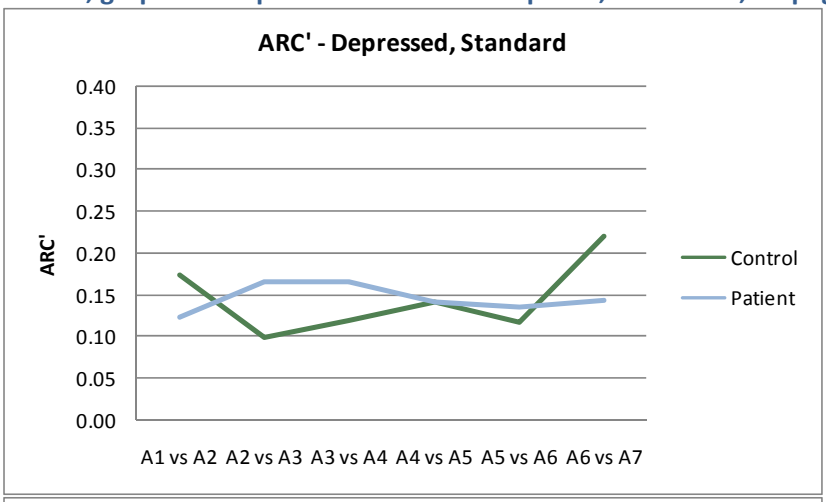


Figure A5.10: Recall on trails 1-7 of the verbal learning test for depressed patients who received standard administration; graph accompanies the data in Chapter 6, Table 6.13, on page 148.

A



B

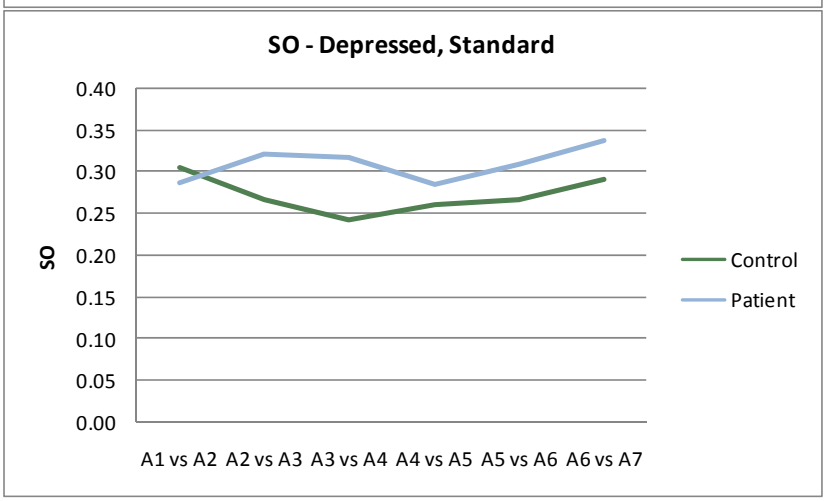


Figure A5.11: A) Subjective organisation as measured by ARC' between recall trial pairs of the verbal learning test for euthymic patients with standard administration; B) Data for the SO measure; graph accompanies data in Chapter 6, Table 6.14, on page 149.

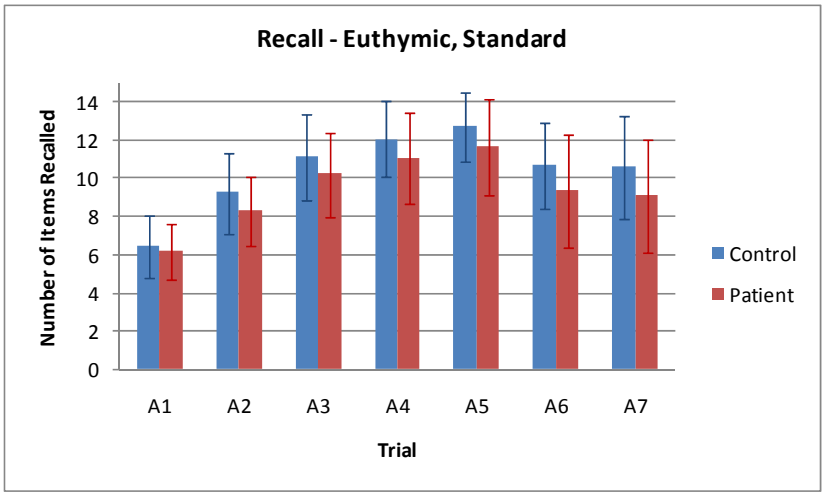
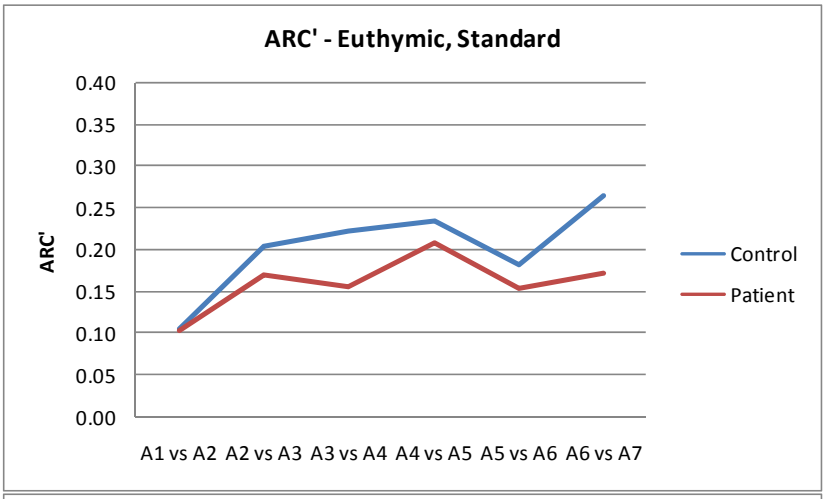


Figure A5.12: Recall on trails 1-7 of the verbal learning test for euthymic patients who received standard administration; graph accompanies the data in Chapter 6, Table 6.18, on page 153.

A



B

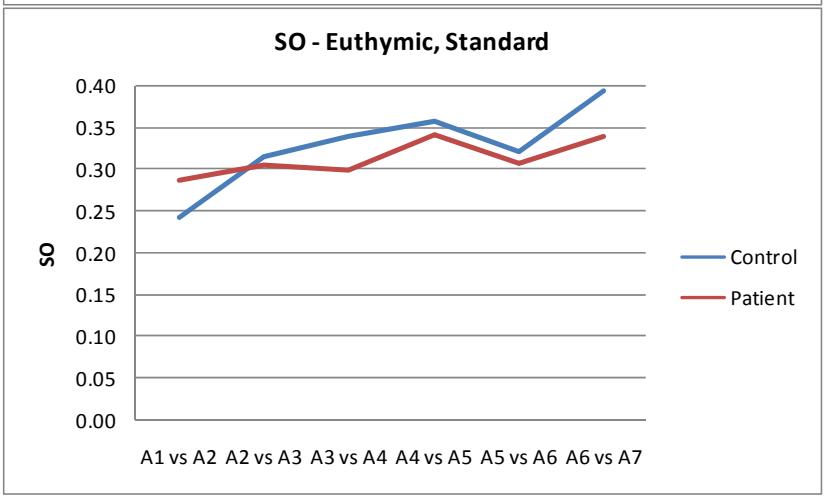


Figure A5.13: A) Subjective organisation as measured by ARC' between recall trial pairs of the verbal learning test for euthymic patients with standard administration; B) Data for the SO measure; graph accompanies data in Chapter 6, Table 6.19, on page 155

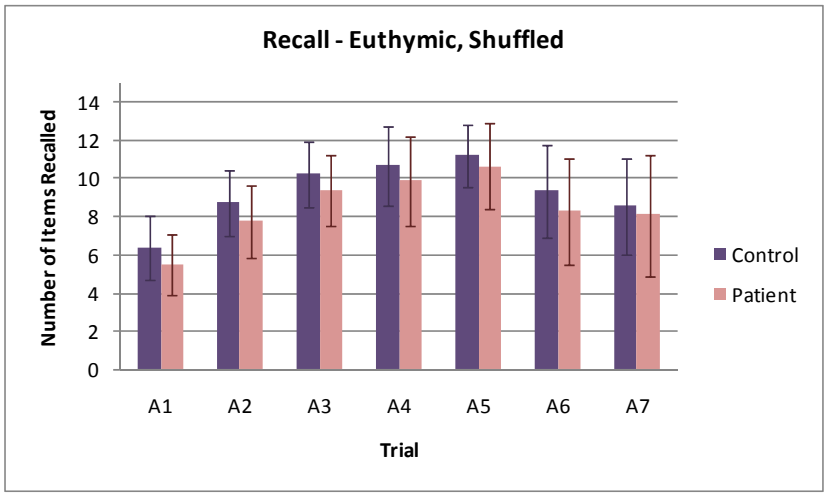


Figure A5.14: Recall on trails 1-7 of the verbal learning test for euthymic patients who received standard administration; graph accompanies the data in Chapter 6, Table 6.23, on page 160.

A

B

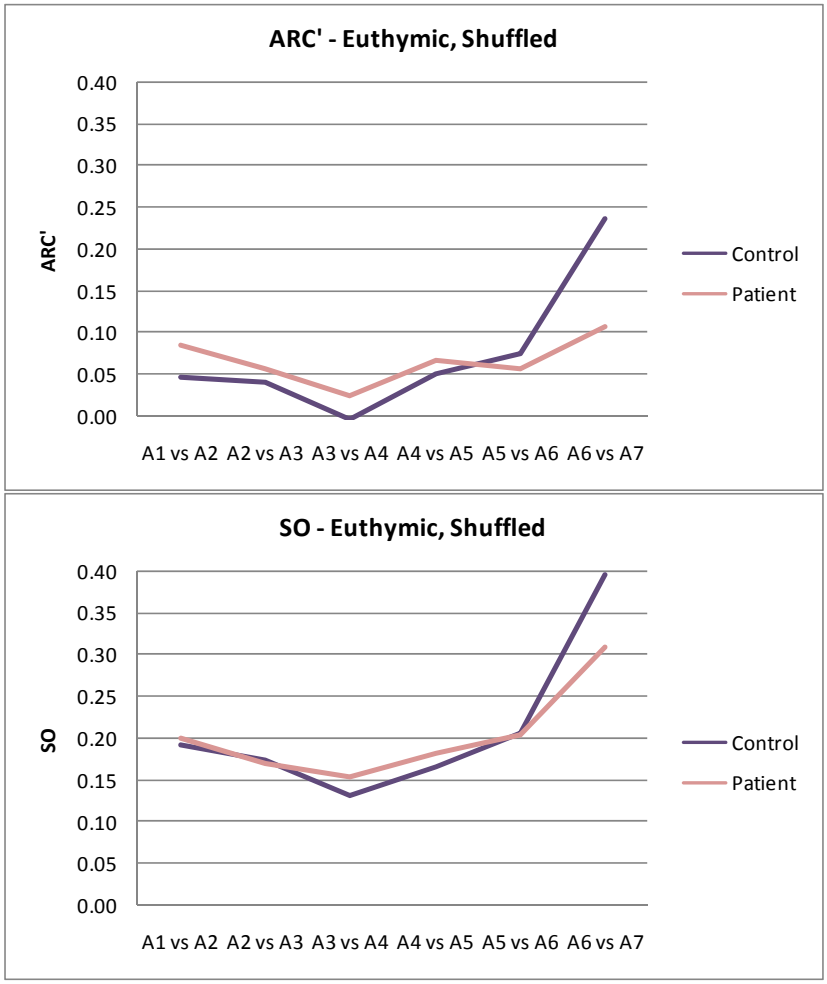


Figure A5.15: A) Subjective organisation as measured by ARC' between recall trial pairs of the verbal learning test for euthymic patients with standard administration; B) Data for the SO measure; graph accompanies data in Chapter 6, Table 6.24, on page 161.

APPENDIX 6: CORRELATION MATRICES

Table A6.53: Spearman correlations between functioning measures and illness history variables (p-values are in *grey italics*)

	GAF	LIFE RIFT				SAS-SR		
		Work	Interpersonal	Recreation	Total	Work	Social & Leisure	Total
Age at onset of first mood episode	0.10 *0.56*	-0.23 *0.18*	0.10 *0.57*	0.10 *0.58*	0.01 *0.94*	-0.01 *0.95*	0.28 *0.10*	0.18 *0.30*
Age at diagnosis	-0.04 *0.84*	-0.15 *0.40*	-0.03 *0.87*	0.24 *0.16*	0.00 *0.98*	-0.13 *0.44*	0.13 *0.45*	-0.10 *0.58*
Number of mood episodes in lifetime	-0.10 *0.58*	-0.13 *0.45*	-0.03 *0.87*	-0.09 *0.63*	-0.01 *0.94*	0.02 *0.91*	-0.04 *0.82*	0.08 *0.66*
Number of previous hospitalisations	-0.20 *0.31*	-0.05 *0.80*	0.07 *0.74*	-0.08 *0.67*	0.03 *0.87*	0.24 *0.21*	0.18 *0.36*	0.21 *0.26*

Table A6.54: Pearson correlations between functional measures (p-values are in *grey italics*)

		GAF	LIFE-RIFT				SAS-SR	
			Work	Interpersonal	Recreation	Total	Work	Social & Leisure
LIFE-RIFT	**Work**	-0.73 *<0.001*						
	Interpersonal	-0.46 *<0.001*	0.45 *<0.001*					
	Recreation	-0.45 *<0.001*	0.46 *<0.001*	0.27 *0.038*				
	Total	-0.75 *<0.001*	0.85 *<0.001*	0.72 *<0.001*	0.69 *<0.001*			
SAS-SR	**Work**	-0.68 *<0.001*	0.56 *<0.001*	0.41 *0.001*	0.31 *0.017*	0.59 *<0.001*		
	Social & Leisure	-0.47 *<0.001*	0.34 *0.007*	0.50 *<0.001*	0.36 *0.005*	0.53 *<0.001*	0.53 *<0.001*	
	Total	-0.59 *<0.001*	0.47 *<0.001*	0.45 *<0.001*	0.40 *0.002*	0.59 *<0.001*	0.77 *<0.001*	0.83 *<0.001*

Table A6.55: Pearson correlations between the factor scores derived from the factor analyses (described in Chapter 9 on page 238; p-values are in *grey italics*)

	Executive	Memory	Speeded processing	Stroop CW	Psychological	Disengagment	Engagement	Emotion Identification
Memory	0.00 *1.00*							
Speeded Processing	0.00 *1.00*	0.00 *1.00*						
Stroop CW	0.00 *1.00*	0.00 *1.00*	0.00 *0.99*					
Psychological	0.14 *0.28*	-0.03 *0.78*	0.24 *0.05*	0.12 *0.32*				
Disengagment	-0.03 *0.80*	0.02 *0.89*	0.04 *0.76*	0.00 *0.99*	0.04 *0.76*			
Engagement	0.06 *0.65*	0.04 *0.76*	0.30 *0.01*	0.12 *0.34*	0.10 *0.40*	0.00 *1.00*		
Emotion Identification	-0.46 *<0.001*	0.19 *0.12*	-0.18 *0.14*	-0.18 *0.15*	-0.12 *0.34*	0.00 *1.00*	0.00 *1.00*	
Semantic Stroop	0.11 *0.40*	-0.07 *0.60*	-0.08 *0.51*	-0.07 *0.55*	0.10 *0.44*	0.00 *1.00*	0.00 *1.00*	0.00 *1.00*

APPENDIX 7: PSYCHOSOCIAL FUNCTIONING SCALES

Clinician-rated measures:

1. Global Assessment of functioning (GAF) (page 325)
2. The Longitudinal Interval Follow-up Evaluation Range of Impaired Functioning Tool (LIFE-RIFT) (page 326)
3. The Specific Levels of Functioning (SLOF) (page 329)

Self-rated measures:

4. Medical Outcomes Survey Short Form 36-item (SF-36) (page 332)
5. The Life Functioning Questionnaire (LFQ) (page 333)

Please note, the formatting of the questionnaires is not as was presented to the participants (no questions ran across pages for example). The SAS-SR cannot be reproduced here for copyright reasons.

DSM-IV Axis V: Global Assessment of Functioning Scale

Consider psychological, social, and occupational functioning on a hypothetical continuum of mental health-illness. Do not include impairment in functioning as a result of physical (or environmental) limitations.

CODE (Note: Use intermediate codes when appropriate, e.g., 45, 68, 72)

GAF rating:
current: _____
highest
past year: _____

100	**Superior functioning in a wide range of activities, life's problems never seem to get out of hand, is sought out by others because of his or her many positive qualities. No symptoms.**
91	**Absent or minimal symptoms** (e.g., mild anxiety before an exam)**; good functioning in all areas, interested and involved in a wide range of activities, socially effective, generally satisfied with life, no more than everyday problems or concerns** (e.g., an occasional argument with family members)**.**
81	**If symptoms are present, they are transient and expectable reactions to psychosocial stressors** (e.g. difficulty concentrating after family argument)**; no more than slight impairment in social, occupational, or school functioning** (e.g., temporarily falling behind in schoolwork).
70	**Some mild symptoms** (e.g., depressed mood and mild insomnia**) OR some difficulty in social, occupational, or school functioning** (e.g., occasional truancy, or theft within the household), **but generally functioning pretty well, has some meaningful interpersonal relationships.**
61	**Moderate symptoms** (e.g., flat affect and circumstantial speech, occasional panic attacks) **OR moderate difficulty in social, occupational, or school functioning** (e.g., few friends, conflicts with peers or co-workers).
50	**Serious symptoms** (e.g., suicidal ideation, severe obsessional rituals, frequent shoplifting) **OR any serious impairment in social, occupational or school functioning** (e.g., no friends unable to keep a job).
40	**Some impairment in reality testing or communication** (e.g., speech is at times illogical, obscure, or irrelevant) **OR major impairment in several areas, such as work or school, family relations, judgment, thinking, or mood** (e.g., depressed man avoids friends, neglects family, and is unable to work; child frequently beats up younger children, is defiant at home, and is failing at school).
31	**Behaviour is considerably influenced by delusions or hallucinations OR serious impairment in communication or judgement** (e.g., sometimes incoherent, acts grossly inappropriately, suicidal preoccupation) **OR inability to function in almost all areas** (e.g., stays in bed all day; no job, home, or friends).
20	**Some danger of hurting self or others** (e.g., suicide attempts without clear expectation or death, frequently violent, manic excitement**) OR occasionally fails to maintain minimal personal hygiene** (e.g., smears faeces**) OR gross impairment in communication** (e.g., largely incoherent or mute).
10	**Persistent danger of severely hurting self or others** (e.g., recurrent violence) **OR persistent inability to maintain minimal personal hygiene OR serious suicidal act with clear expectation of death.**
0	Inadequate information

The LIFE-RIFT

Code:

(1a) Employment: Which of the following categories best characterizes the degree to which the patient's current (past week) work activities have been impaired as a result of psychopathology?

☐ 0 **Not applicable**. Did not work during the past week, for reasons other than psychopathology.

☐ 1 **No impairment – high level**. Worked as much as someone in his social situation would be expected to work, and worked at a high level.

☐ 2 **No impairment – satisfactory level**. Worked as much as someone in his social situation would be expected to work, and worked at a satisfactory level.

☐ 3 **Mild impairment**. Worked somewhat less that someone in his social situation would be expected work and/or had mild difficulties in carrying out work activities.

☐ 4 **Moderate impairment**. Has missed a lot of work and/or has had considerable difficulties in carrying out work activities.

☐ 5 **Severe impairment**. Has missed a great deal of work when someone in his social situation would have been expected to work and/or has been virtually unable to carry out his work activities when he did work.

☐ 6 **No information**.

(1b) Household: Which of the following categories best characterizes the degree to which he patient's current (past week) household work has been impaired as a result of psychopathology?

☐ 0 **Not applicable**. Did not carry out household duties during the past week, for reasons other than psychopathology.

☐ 1 **No impairment – high level**. Has carried out housework most of the time that would be expected, and worked at a high level.

☐ 2 **No impairment – satisfactory level**. Has carried out housework most of the time that would be expected and worked at a satisfactory level

☐ 3 **Mild impairment**. Worked somewhat less than expected and/or had mild difficulties in carrying out housework.

☐ 4 **Moderate impairment**. Has missed a lot of housework when expected and/or has had considerable difficulties in carrying out housework.

☐ 5 **Severe impairment**. Has missed a great deal of housework when expected to work and/or has been virtually unable to carry out housework when he attempts it.

☐ 6 **No information**.

(1c) Student: _____

Which of the following categories best characterizes the degree to which the patient's current school work has been impaired as a result of a psychopathology?

☐ 0 **Not applicable**. Because not currently enrolled in a student programme for reasons other than psychopathology.

☐ 1 **No impairment – high level**. Worked as much as would be expected if not symptomatic and got high grades.

☐ 2 **No impairment – satisfactory level**. Worked as much as would be expected if not symptomatic and got satisfactory grades.

☐ 3 **Mild impairment**. Worked somewhat less and/or got grades below expected if not symptomatic.

☐ 4 **Moderate impairment**. Missed a lot of school work and/or got grades consistently below expected.

☐ 5 **Severe impairment**. Missed most of school work and/or dropped out of school or got grades far below those expected.

☐ 6 **No information**.

(1) Work (maximum of 1a, 1b and 1c) : _____

(2) Interpersonal Relations: Which of the following best characterizes the patient's level of interpersonal relationships with his family currently (past month)? [Provide separate ratings for spouse (2a), children (2b) and other relatives (2c).]

(2a) Interpersonal relations with spouse: _____
(2b) Interpersonal relations with children: _____
(2c) Interpersonal relations with other relatives: _____

0 **Not applicable** because does not have relatives in this category.

1 **Very good**. Experiences very good relationships with this/these family member(s), with only transient friction which is rapidly resolved. Feels only very minor or occasional need to improve quality of relationship, which is usually close and satisfying.
2 **Good**. Argues occasionally, but arguments usually resolve satisfactorily within a short time. May occasionally prefer not to be with then because of dissatisfaction with them or be actively working with them to improve relationship.
3 **Fair**. Often argues with this (these) family member(s) and takes a long time to resolve arguments. May withdraw from this person (these people) due to dissatisfaction. Often thinks that relationship needs to be either more harmonious or closer emotionally even when no conflict is present. For those relatives not living with the subject, contacts with them by choice are less frequent than feasible or rarely enjoyed very much when made.
4 **Poor**. Regularly argues with this (these) family member(s) and such arguments are rarely ever resolved satisfactorily. Regularly prefers to avoid contact with them and/or feels great deficit in emotional closeness. For those family members out of the household, subject avoids seeing them as much as possible and derives no pleasure from contact when made.
5 **Very poor**. Either constantly argues with this (these) family member(s) or withdraws from them most of the time. Separated or divorced from spouse or children moved out of household or almost always hostile to them when in contact.
6 **Variable**. Different levels for various members of this group, and would warrant a rating of good or better (2, 1) with at least 1 member of this group (rate as 2)
7 **Variable**. Different levels for various members of this group, and would not warrant a rating of good or better (2, 1) with any member of his group. (Rate as 4).
8 **No information**.

(2d) Interpersonal relations with friends: Which of the following best characterizes the patient's interpersonal relationships with friends currently (past month)?

☐ 1 **Very good**. Had several special friends that he saw regularly and frequently was close to.
☐ 2 **Good**. Had at least two special friends that he was from time to time and was fairly close to.
☐ 3 **Fair**. Had only one special friend that he saw from time to time and was fairly close to; or contacts very limited to several friends that he was not very close to emotionally.
☐ 4 **Poor**. Had no special friends he saw from time to time and was fairly close to; or contacts limited to one or two friends that he was not very close to.
☐ 5 **Very poor**. Had no special friends and practically no social contacts.

☐ 6 **No information**.

(2) Interpersonal relations (maximum of 2a, 2b, 2c and 2d): _____

(3) Satisfaction: Which of the following best characterizes the patient's overall level of satisfaction (contentment, degree to which he feels fulfilled, gratification derived from activities) for the past week?:

☐ 1 **Very good**. Transient problems may occur, but generally satisfied with all aspects of his life. Occasional minor dissatisfaction in one area, but overall is quite content with himself, job, family, friends, activities and finances.

☐ 2 **Good**. Mild dissatisfaction persists, but only in one area or is intermittent in several areas. In balance, is generally content and able to enjoy life most of the time, but does think there should be some improvement in either occupational role, interpersonal relations, sexual activities or finances.

☐ 3 **Fair**. Moderate dissatisfaction in one or more areas, which is relatively persistent. Either discontent with occupational role, interpersonal relations, sexual activities or finances.

☐ 4 **Poor**. Very dissatisfied in most areas and derives little pressure from life. Rarely able to derive any satisfaction from activities or relationships.

☐ 5 **Very poor**. Derives no satisfaction from anything. May feel no desire to carry out the smallest task or to be with other people.

☐ 6 **No information**.

(4) Recreation: At what level has the patient been involved in and able to enjoy recreational activities and hobbies (reading, spectator or participant sports, gardening, music, sewing, attending parties or gatherings, church or community organisations) in the past week.

☐ 1 **Very good**. Has at least two activities which he enjoys fully and frequently.

☐ 2 **Good**. Participates in several activities and does not always fully enjoy them; or participates in fewer activities or less frequently that optimal but enjoys participation.

☐ 3 **Fair**. Occasional participation in recreational activities or hobbies; or limited enjoyment when participation occurs.

☐ 4 **Poor**. Some participation in recreational activities or hobbies, and derives very little enjoyment from such activities.

☐ 5 **Very poor**. No involvement in recreational activities or hobbies.

☐ 6 **No information**.

SUMMARY

(1) Work (maximum of 1a, 1b and 1c): _____
(2) Interpersonal relations (maximum of 2a, 2b, 2c and 2d): _____
(3) Satisfaction: _____
(4) Recreation: _____

Total score (sum of 1, 2, 3 and 4):

Taken from: Leon, AC, Solomon, DA, Mueller, TI, Turvey, CL, Endicott, J & Keller, MB (1999) The Range of Impaired Functioning Tool (LIFE-RIFT): a brief measure of functional impairment. *Psychological Medicine* 29(4), pp. 869-78.

Specific Levels of Functioning

Instructions: Circle the number that best describes this person's *typical* level of functioning on each item listed below. *Be as accurate as you can*. If you are not sure about a certain rating, ask someone who may know or consult the case record.
Mark only one number for each item. Be sure to mark all items.

Self-Maintenance

A. Physical Functioning	No Problem	Problem, but no effect on General Functioning	Slight Effect on General Functioning	Restricts General Functioning Substantially	Prevents General Functioning
1. Vision	5	4	3	2	1
2. Hearing	5	4	3	2	1
3. Speech impairment	5	4	3	2	1
4. Walking, use of legs	5	4	3	2	1
5. Use of hands and arms	5	4	3	2	1

B. Personal Care Skills	Totally Self-sufficient	Needs Verbal Advice or Guidance	Needs Some Physical Help or Assistance	Needs Substantial Help	Totally Dependent
6. Toileting (uses toilet properly; keeps self and area clean)	5	4	3	2	1
7. Eating (uses utensils properly; eating habits)	5	4	3	2	1
8. Personal hygiene (body and teeth; general cleanliness)	5	4	3	2	1
9. Dressing self (selects appropriate garments; dresses self)	5	4	3	2	1
10. Grooming (hair, make-up, general appearance)	5	4	3	2	1
11. Care of own possessions	5	4	3	2	1
12. Care of own living space	5	4	3	2	1

Social Functioning

C. Interpersonal Relationships	Highly Typical of This Person	Generally Typical of This Person	Somewhat Typical of This Person	Generally Untypical of This Person	Highly Untypical of This Person
13. Accepts contact with others (does not withdraw or turn away)	5	4	3	2	1
14. Initiates contact with others	5	4	3	2	1
15. Communicates effectively (speech and gestures are understandable and to the point)	5	4	3	2	1
16. Engages in activities without prompting	5	4	3	2	1
17. Participates in groups	5	4	3	2	1
18. Forms and maintains friendships	5	4	3	2	1
19. Asks for help when needed	5	4	3	2	1

D. Social Acceptability	Never	Rarely	Sometimes	Frequently	Always
20. Verbally abuses others	5	4	3	2	1
21. Physically abuses others	5	4	3	2	1
22. Destroys property	5	4	3	2	1
23. Physically abuses self	5	4	3	2	1
24. Is fearful, crying, clinging	5	4	3	2	1
25. Takes property from others without permission	5	4	3	2	1

26. Performs repetitive behaviours (pacing, rocking, making noises)	5	4	3	2	1

Community Living Skills

E. Activities	Totally Self-sufficient	Needs Verbal Advice or Guidance	Needs Some Physical Help or Assistance	Needs Substantial Help	Totally Dependent
27. Household responsibilities (house cleaning, cooking, washing clothes)	5	4	3	2	1
28. Shopping (selection of items, choice of shops, payment at till)	5	4	3	2	1
29. Handling personal finances (budgeting, paying bills)	5	4	3	2	1
30. Use of telephone (getting number, dialling, speaking, listening)	5	4	3	2	1
31. Travelling from residence without getting lost	5	4	3	2	1
32. Use of public transportation (selecting route, using timetable, paying fares, making transfers)	5	4	3	2	1
33. Use of leisure time (reading, visiting friends, listening to music)	5	4	3	2	1
34. Recognising and avoiding common dangers (traffic safety, fire safety)	5	4	3	2	1
35. Self-medications (understanding purpose, taking as prescribed, recognising side effects)	5	4	3	2	1
36. Use of medical and other community services (knowing whom to contact, how, and when to use)	5	4	3	2	1
37. Basic reading, writing, and arithmetic (enough for daily needs)	5	4	3	2	1

F. Work Skills	Highly Typical of This Person	Generally Typical of This Person	Somewhat Typical of This Person	Generally Untypical of This Person	Highly Untypical of This Person
38. Has employable skills	5	4	3	2	1
39. Works with minimal supervision	5	4	3	2	1
40. Is able to sustain work effort (not easily distracted, can work under stress)	5	4	3	2	1
41. Appears at appointments on time	5	4	3	2	1
42. Follows verbal instructions accurately	5	4	3	2	1
43. Completes assigned tasks	5	4	3	2	1

Other Information

44. From your knowledge of this person, are there other skills or problem areas not covered on this form that are important to this person's ability to function independently? If so, please specify:

45. How well do you know the skills and behaviour or the person you just rated? (circle one):

Very Well		Fairly Well		Not Very Well At All
1	2	3	4	5

46. Have you discussed this assessment with the client? (circle one) Yes No

If yes, does the client generally agree with the assessment? (circle one) Yes No

SF36 Health Survey

INSTRUCTIONS: This set of questions asks for your views about your health. This information will help keep track of how you feel and how well you are able to do your usual activities. Answer every question by marking the answer as indicated. If you are unsure about how to answer a question please give the best answer you can.

1. In general, would you say your health is: (Please tick one box.)
   - Excellent ☐
   - Very Good ☐
   - Good ☐
   - Fair ☐
   - Poor ☐

2. Compared to one year ago, how would you rate your health in general now? (Please tick one box.)
   - Much better than one year ago ☐
   - Somewhat better now than one year ago ☐
   - About the same as one year ago ☐
   - Somewhat worse now than one year ago ☐
   - Much worse now than one year ago ☐

3. The following questions are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much? **(Please circle one number on each line.)**

	Activities	Yes, Limited A Lot	Yes, Limited A Little	Not Limited At All
3(a)	**Vigorous activities**, such as running, lifting heavy objects, participating in strenuous sports	1	2	3
3(b)	**Moderate activities**, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf	1	2	3
3(c)	Lifting or carrying groceries	1	2	3
3(d)	Climbing several flights of stairs	1	2	3
3(e)	Climbing one flight of stairs	1	2	3
3(f)	Bending, kneeling, or stooping	1	2	3
3(g)	Waling more than a mile	1	2	3
3(h)	Walking several blocks	1	2	3
3(i)	Walking one block	1	2	3
3(j)	Bathing or dressing yourself	1	2	3

4. During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

	(Please circle one number on each line.)	Yes	No
4(a)	Cut down on the **amount of time** you spent on work or other activities	1	2
4(b)	Accomplished less than you would like	1	2
4(c)	Were **limited in the kind** of work or other activities	1	2
4(d)	Had **difficulty** performing the work or other activities (for example, it took extra effort)	1	2

5. During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (e.g. feeling depressed or anxious)?

	(Please circle one number on each line.)	Yes	No
5(a)	Cut down on the amount of time you spent on work or other activities	1	2
5(b)	Accomplished less than you would like	1	2
5(c)	Didn't do work or other activities as **carefully** as usual	1	2

Life Functioning Questionnaire

How much difficulty have you had in the following areas over the past month? (*Tick the box that best describes your **degree of difficulty functioning**, if any, over the past month.)*

1. Leisure Time

A. Leisure activities with friends

(If you never spend time with your friends, or if you do not have any friends, please tick this box ☐ and go to '**B**')

	Degree of Difficulty Functioning			
	No problems	**Mild problems**	**Moderate problems**	**Severe problems**
1. Time: amount of time spent with friends	☐	☐	☐	☐
2. Conflict: getting along with friends	☐	☐	☐	☐
3. Enjoyment: enjoying time spent with friends	☐	☐	☐	☐

If you are having ANY difficulty, what do you think is the cause? _____

_____

_____

B. Leisure activities with family

(If you never spend time with your family, or if you have no family, please tick this box ☐ and go to '**C**')

	No problems	**Mild problems**	**Moderate problems**	**Severe problems**
4. Time: amount of time spent with family	☐	☐	☐	☐
5. Conflict: getting along with family	☐	☐	☐	☐
6. Enjoyment: enjoying and having an interest in family activities	☐	☐	☐	☐

If you are having ANY difficulty, what do you think is the cause? _____

_____

_____

2. Duties/Responsibilities

C. Duties at home (e.g. housework, paying bills, grocery shopping, mowing lawn, childcare tasks, car repairs etc.)
(If you have no duties at home, or are homeless, please tick this box ☐ and go to '**D**')

	Degree of Difficulty Functioning			
	No problems	**Mild problems**	**Moderate problems**	**Severe problems**
7. Time: amount of time spent performing duties	☐	☐	☐	☐
8. Conflict: can you perform these duties without undue friction with others?	☐	☐	☐	☐
9. Enjoyment: enjoying time spent with friends	☐	☐	☐	☐
10. Performance: quality of work (doing a good job; getting the job done)	☐	☐	☐	☐

If you are having ANY difficulty, what do you think is the cause? _____

_____

_____

C. Duties at work, school or activity centre
(If you are not working or not in school, please tick this box ☐ and go to next page)

	No problems	**Mild problems**	**Moderate problems**	**Severe problems**
7. Time: amount of time spent at work, school, etc.	☐	☐	☐	☐
8. Conflict: getting along with coworkers and supervisors	☐	☐	☐	☐
9. Enjoyment: enjoyment/satisfaction and interest from work	☐	☐	☐	☐
10. Performance: quality of work	☐	☐	☐	☐

If you are having ANY difficulty, what do you think is the cause? _____

_____

_____

How many days did you miss over this **last month** at work or school due to your mental illness?

Work
1. ☐ Not applicable
2. ☐ 0-5 days
3. ☐ 6-10 days

School
1. ☐ Not applicable
2. ☐ 0-5 days
3. ☐ 6-10 days

4. ☐ 11-20 days
5. ☐ over 20 days

4. ☐ 11-20 days
5. ☐ over 20 days

D. Reasons causing difficulty in functioning

Did any of the factors below cause you difficulties at work this month, or cause you to work less than full-time, or not at all? (*Please mark all that apply for this month*.)

1. ☐ Too depressed most of the time
2. ☐ Too manic most of the time
3. ☐ Couldn't get my mood stable long enough to work – too up and down
4. ☐ Afraid to work at usual level because afraid of precipitating another episode
5. ☐ Wanted to work but the kind of job I could get due to the gaps in my work history was too demeaning for my educational level
6. ☐ Mood OK and wanted to work but couldn't get a job due to the gaps in my work history
7. ☐ Couldn't get along with others
8. ☐ Wanted my old job but couldn't get it
9. ☐ Could get my old job but felt embarrassed to go back
10. ☐ Disability cheque was greater than I could have earned otherwise
11. ☐ Didn't have a job for a long time prior to my most recent episode
12. ☐ Physical symptoms (e.g. difficulty concentrating, blurred vision, fatigue/sedation) interfered with my functioning
13. ☐ Didn't need to work (e.g. retired, supported by someone else, etc) but could if need be
14. ☐ Medication side effects interfered with functioning
15. ☐ Other (*please explain*):

E. Please tick the box of the answer(s) which best describes your situation:

1. Work situation this month (please tick only those boxes that apply in the last *30 days*):

- **Competitive job** (paid job obtained without assistance of rehab programme)
  - ☐ 1. Full-time at same or higher job level than that held prior to most recent episode
  - ☐ 2. Part-time at same or higher job level than that held prior to most recent episode
  - ☐ 3. Full-time at lower job level than that held prior to most recent episode
  - ☐ 4. Part-time at lower job level than that held prior to most recent episode
- **Transitional job** (paid job obtained through vocational rehabilitation programme)
  - ☐ 5. Full-time
  - ☐ 6. Part-time
- **Work Training**
  - ☐ 7. Work training
- **Sheltered Workshop**
  - ☐ 8. Sheltered workshop
- **Volunteer**
  - ☐ 9. Full-time
  - ☐ 10. Part-time
- **Student**
  - ☐ 11. Full-time
  - ☐ 12. Part-time
- **Housewife/husband**
  - ☐ 13. As full-time job
  - ☐ 14. As part-time job
- **Not working in job, school or home**
  - ☐ 15. Not working in job, school or home
- **Other**
  - ☐ 16. Other (*please explain*): _____

_____

2. How many days per week are you scheduled to attend:

1. Work _____ 2. School _____

3. Day Hospital _____ 4. Activity Centre _____

3. Living situation over the last six months (*please tick all that apply*):

1. ☐ Hospital
2. ☐ Skilled nursing facility – 24-hour nursing service
3. ☐ Intermediate care facility – less than 24-hour nursing care facility
4. ☐ Supervised group living (long-term)
5. ☐ Transitional group home (halfway or quarterway house)
6. ☐ Family foster care
7. ☐ Cooperative apartment, supervised (staff on premises)
8. ☐ Cooperative apartment, unsupervised (staff no on premises)
9. ☐ Board and care home (private proprietary home for adults, with programme supervision
10. ☐ Boarding house (includes meals, no programme or supervision)
11. ☐ Rooming or boarding house or hotel (includes single room occupancy, no meals are provided, cooking facilities may be available)
12. ☐ Private house or apartment
13. ☐ Shelter
14. ☐ Jail
15. ☐ No residence (that is, you often need to live/sleep on the streets, or other areas not generally intended for residence)

4. Financial situation over last six months (*please tick all that apply*):

1. ☐ Received no pay (fully supported by someone else; e.g. parents, spouse)
2. ☐ Received wages for work performed
3. ☐ Received benefits
4. ☐ Received retirement benefits or pension
5. ☐ Other (*please specify*): _____

_____

5. A) When did you last work full-time? (*Please tick only ONE box*)

1. ☐ I work full-time now (YOU HAVE FINISHED THE QUESTIONNAIRE)
2. ☐ I have never worked full-time
3. ☐ Within the last 2 years
4. ☐ 2-5 years ago
5. ☐ 5-10 years ago
6. ☐ Over 10 years ago

B) How long were you working full-time the last time you worked full-time? (*Please only tick ONE box*):

1. ☐ Less than one month
2. ☐ Less than 6 months
3. ☐ Less than 1 year
4. ☐ 1 year or more

C) Why did you stop working full-time? (If more than one reason, please rank in order of importance: 1=most important, 2=next most important, etc.)

Ranking			
_____	1.	☐	Mental illness
_____	2.	☐	Physical illness
_____	3.	☐	Children
_____	4.	☐	Couldn't find job after leaving/being laid off from previous job
_____	5.	☐	Retired
_____	6.	☐	Other (*please explain*): _____

_____

This is the end of the questionnaire.

Thank you.

Taken from: Altshuler, Mintz & Leight (2002) The Life Functioning Questionnaire (LFQ): a brief, gender-neutral scale assessing functional outcome
Journal of Affective Disorders, 112, pp. 161-182