Spatial Memory Processes in Bipolar Depression:

neuropsychological and HPA axis correlates

Peter Gallagher

A thesis submitted in partial fulfilment of the requirements of the regulations for the degree of Doctor of Philosophy

Newcastle University
Faculty of Medical Sciences
Institute of Neuroscience

Table of Contents

List of tables

List of figures

Acknowledgements

Firstly, I would like to give my thanks to my supervisors – to Dr John Gray, for his invaluable guidance in navigating through the experimental neuropsychology maze; to Professor Allan Young, for first introducing me to the world of mood disorders research and supporting my research through the years, even during his transatlantic sojourn; and to Professor Nicol Ferrier, for agreeing to take on the baton and for expanding my research horizons. I thank each of them for the unique expertise they brought to the project and to this thesis.

Thank you to Dr Roy Kessels and Dr Colin Hamilton for their input and helpful discussions around the visuospatial tests and processes examined.

I am eternally grateful to Dr Stuart Watson and an excellent Research Nurse, Mrs Margaret Smith, who helped in so many different ways with the design, running and completion of these studies from start to finish. I also thank many others who gave their time and contributed to different aspects of recruitment, data collection and technical support: Ms Anna Massey, Ms Lisa Svennson, Mrs Helen Watkinson, Mr Mel Leitch, Mrs Shirley Dodds, Dr Jane Carlile, Dr Dolores Del Estal, Dr Andrea Hearn, Dr Samer Makhoul, Dr Anu Menon, Dr Bruce Owen, Dr Harikumar Ramachandran.

I would also like to thank the Stanley Medical Research Institute, Medical Research Council and the Mental Health Foundation for funding the program of research of which these studies are part.

On a personal level, I would like to thank many friends and family: my parents, sister and grandparents for their encouragement to head off in this direction in life and their unwavering support ever since; to Lucy Robinson, for her endless support and motivation to complete; and to Jenny Nendick, Alex Robinson, Gill Brown, Jill Thompson and Paul Mackin for helping me relax over a good pint or two.

Finally, I would first like to that all the participants who gave up their time to take part in these studies. Without them we literally have nothing.

Abstract

Background/ aims: Bipolar disorder is associated with significant impairment in a broad range of neuropsychological processes in addition to hypothalamic-pituitary-adrenal (HPA) axis dysfunction and hypercortisolaemia. As both animal and human models have highlighted the role of cortisol in the modulation of memory processes, attempting to understand this link is of critical importance. The aims of this thesis are to first profile neuropsychological and HPA axis function in individuals with a diagnosis of bipolar disorder, before examining if these functions can be altered through an intervention with an antiglucocorticoid drug. The subsequent chapters of this thesis will report analyses designed to explore specific aspects of these changes in more detail, principally alterations in spatial memory processes.

Method: The thesis reports two broad phases of research. The first is a study of 20 participants diagnosed with bipolar disorder (with depressive symptoms) who first completed a broad neuropsychological assessment and profiling of afternoon cortisol and DHEA levels. These individuals then entered a randomised crossover study to examine the effects of mifepristone (RU-486), a glucocorticoid receptor antagonist, on neuropsychological functions and mood. A second cohort of 53 participants diagnosed with bipolar depression (BD) and 47 healthy controls was recruited to explore aspects of the results in more detail, particularly the fractionation of spatial memory and the integration of neuropsychological processes and their relationship with measures of HPA axis function.

Results: 1) BD participants exhibited broad neuropsychological impairment across a range of cognitive domains in addition to hypercortisolaemia. 2) Administration of an antiglucocorticoid drug significantly reduced cortisol levels and improved spatial working memory performance. 3) The underlying neuropsychological component structure of BD and controls differed. 4) BD participants exhibited impairments in fine-grain metric spatial memory which, unlike other spatial processes, could not be explained by other measures. 5) A unique profile of processes underpinning aspects of visuospatial memory was observed in BD, suggesting a form of cognitive 'scaffolding'. 6) A simple link between neuropsychological processes and peripheral HPA axis measures was not observed.

Conclusion: Spatial memory processes in bipolar depression can be altered by direct HPA axis manipulation. A number of interesting avenues for future research have been identified that will further our knowledge of the integration between the biological mechanisms underlying neuropsychological impairment in mood disorders and should develop our understanding of integration between cognitive processes in general.

Chapter I

A General Introduction

1. General Introduction

1.1 Outline and organization of the thesis

The following thesis contains seven chapters: a general introduction, five empirical chapters, and a general discussion.

Chapter one, the general introduction, presents an overview of bipolar disorder and a focussed review of neuropsychological and hypothalamic-pituitary-adrenal (HPA) axis abnormalities within the disorder. This provides the background to chapters two and three which contain data from the same cohort of patients from a program of research examining the effectiveness of antiglucocorticoid treatment in bipolar disorder. Chapter two examines the broad neuropsychological performance and peripheral cortisol levels in this group, while chapter three reports the effects of an antiglucocorticoid (mifepristone) on these measures. A discussion of these findings highlights several important areas for further study and identifies methodological improvements.

The subsequent three empirical Chapters (Chapters four to six) are born out of the findings of the initial studies and explore in more depth specific hypotheses resulting from these. Chapter four examines the factor structure of neuropsychological measures in patients with bipolar disorder in comparison to that in healthy controls. Chapter five reports the findings of a more specific analysis of – and fractionation of – spatial memory processes in bipolar depression. Finally, as a result of the findings of Chapters four and five, Chapter six examines the relationship between specific spatial memory processes and broader neuropsychological factors and the relationship with other HPA axis measures. Chapter seven contains the general discussion of all of these findings.

1.2 Bipolar disorder: an overview

In this section, a brief overview of bipolar disorder is presented, covering the historical conceptualisation, current diagnostic classification, clinical features and epidemiology.

1.2.1 History

The origins of the concept of bipolar disorder can be traced back to the classical period and are found in the writings of Greek physicians and philosophers. Hippocrates (460–337 BC) was the first to systematically describe mania and melancholia, supported by clinical observation and extended follow-up (Angst & Marneros, 2001). His theories were progressed by many scholars, such as Aretaeus of Cappadocia who, in the 1st century AD, published works such as 'On the Aetiology and Symptomatology of Chronic Diseases' and 'The Treatment of Chronic Diseases' in which he described mental disorders in detail and was the first to link mania and melancholia (Angst & Marneros, 2001).

Over the centuries many others explored and developed these ideas. The explicit conceptualisation of BD as a single distinct entity emerged in the 19th century with the work of Jean-Pierre Falret. In 1851 Falret published a statement in the Gazette des Hôpitaux in which he described a single disorder which he named *'folie circulaire'*, characterized by a continuous cycle of depression, mania and 'well' intervals of varying length (Falret, 1851; Angst & Marneros, 2001). This notion gained popularity throughout Europe (although not without some opposition). However, it is undoubtedly the work of the German psychiatrist Emil Kraepelin which laid the foundation for the concept of bipolar disorder as we know it today.

Kraepelin was the first to divide psychosis into two discrete disorders: dementia praecox (schizophrenia) and manic-depressive insanity (bipolar disorder). By 1913, in the eighth edition of Kraepelin's text, virtually all of the major clinical forms of melancholia had been subsumed under "manic-depressive illness" – a state that he argued could be clearly differentiated from dementia praecox by its periodic or episodic course, a more benign prognosis and a family history of manic-depressive illness (Goodwin & Jamison, 2007).

As with other significant periods in the development of the bipolar concept, there were some who did not support this system. For example, Carl Wernicke – and later Karl Kleist – both argued against the notion of unifying many conditions under the bipolar concept, which they saw as distinct, especially melancholia (Wernicke, 1900; Kleist, 1911). However, it has been noted that these detailed distinctions did not gain widespread acceptance because of their complexity (Angst & Marneros, 2001). The foundations laid down by Kraepelin were subsequently taken forward through the 'research diagnostic criteria' era and refined through the subsequent iterations of these to the formal diagnostic systems we have today.

In the following section the current formal classification of bipolar disorder and its clinical features are discussed. While these are utilised in the context of the present study, they are far from objective and they are not uncritically accepted. The current edition of ICD-10 describes the situation thus: "*It seems likely that psychiatrists will continue to disagree about the classification of disorders of mood until methods of dividing the clinical syndromes are developed that rely at least in part upon physiological or biochemical measurement, rather than being limited as at present to clinical descriptions of emotions and behaviour. As long as this limitation persists, one of the major choices lies between a comparatively simple classification with only a few degrees of severity, and one with greater details and more subdivisions…. Options for specifying several aspects of affective disorders have been included*

{in ICD-10}*, which, although still some way from being scientifically respectable, are regarded by psychiatrists in many parts of the world as clinically useful. It is hoped that their inclusion will stimulate further discussion and research into their true clinical value*" (WHO, 1992).

1.2.2 Diagnosis and clinical features

Bipolar disorder is recognised as a heterogeneous condition, typically characterised by extreme fluctuations in mood that cycle between (hypo)manic or mixed states and depressive episodes interspersed with periods of euthymia. Current formal diagnostic systems such as the ICD-10 and DSM-IV contain similar features of mania and depression although the specific diagnostic categories differ slightly.

ICD-10 describes the disorder as being characterized by "*repeated (i.e. at least two) episodes in which mood and activity levels are significantly disturbed, this disturbance consisting on some occasions of an elevation of mood and increased energy and activity (mania or hypomania), and on others of a lowering of mood and decreased energy and activity (depression)*".

These disturbances of mood are specifically characterized – mania is described as a period (for at least 1 week; severe enough to disrupt ordinary work and social activities more or less completely) where mood is elevated "*out of keeping with the individual's circumstances and may vary from carefree joviality to almost uncontrollable excitement. Elation is accompanied by increased energy, resulting in over activity, pressure of speech, and a decreased need for sleep. Normal social inhibitions are lost, attention cannot be sustained, and there is often marked distractibility. Self-esteem is inflated, and grandiose or over-optimistic ideas are freely expressed"* (WHO, 1992). Within ICD-10 a with- or without- psychotic symptoms distinction is also made.

Hypomania is defined as a lesser degree of mania, in which "*abnormalities of mood and behaviour are too persistent and marked to be included under cyclothymia but are not accompanied by hallucinations or delusions"*. There is a persistent mild elevation of mood (for at least several days on end), increased energy and activity, and usually marked feelings of well-being and both physical and mental efficiency. Increased sociability, talkativeness, over-familiarity, increased sexual energy, and a decreased need for sleep are often present but not to the extent that they lead to severe disruption of work or result in social rejection. Irritability, conceit, and boorish behaviour may take the place of the more usual euphoric sociability. Concentration and attention may be impaired, thus diminishing the ability to settle down to work or to relaxation and leisure, but this may not prevent the appearance of interests in quite new ventures and activities, or mild over-spending (WHO, 1992).

As described above, within bipolar disorder depressive episodes also occur. ICD-10 includes a mild/moderate/severe distinction but characterises all as where an individual suffers from depressed mood, loss of interest and enjoyment, and reduced energy leading to increased fatigue and diminished activity. Marked tiredness after only slight effort is common. Other common symptoms are: reduced concentration and attention; reduced self-esteem and self-confidence; ideas of guilt and unworthiness; bleak and pessimistic views of the future; ideas or acts of self-harm or suicide; disturbed sleep; diminished appetite. The lowered mood is noted to vary little from day to day, and is often unresponsive to circumstances, yet may show a characteristic diurnal variation as the day goes on. A duration of at least 2 weeks is typically required for diagnosis, but shorter periods may be reasonable if symptoms are unusually severe and of rapid onset.

The DSM-IV shares many common features with the ICD-10, although includes a more formal distinction between bipolar-I (includes a history of mania) and bipolar-II (includes a history of

hypomania) (The bipolar-II sub-type is included under a sub-heading of 'other bipolar affective disorders' in ICD-10). It is also specified that hypomania should last for at least 4 days (rather than 'several days' in ICD-10) for a diagnosis of bipolar-II disorder and that the overall diagnosis can be made on the basis of a single manic episode. The features of a major depressive episode are very similar to the ICD-10 and "*require five (or more) of the following symptoms to have been present during the same 2-week period (and represent a change from previous functioning); at least one of the symptoms is either (i) or (ii)"*:

> (i) depressed mood most of the day, nearly every day;
>
> (ii) markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day;
>
> (iii) significant weight loss when not dieting or weight gain, or decrease or increase in appetite nearly every day; (iv) insomnia or hypersomnia nearly every day; (v) psychomotor agitation or retardation nearly every day; (vi) fatigue or loss of energy nearly every day; (vii) feelings of worthlessness or excessive or inappropriate guilt nearly every day; (viii) diminished ability to think or concentrate, or indecisiveness, nearly every; (ix) recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide (APA, 1994).

Other specifiers can also be added to these, including presence or absence of psychosis and rapid cycling (4 or more mood episodes in a one-year period). It should also be noted that these formal diagnostic criteria have been through some development and refinement by authors who see such illnesses as a much more detailed, incremental spectrum from bipolar-I: (full-blown mania) through to bipolar-VI (a late-onset, dementia-associated illness) (Akiskal & Pinto, 1999; Ng et al., 2008). These have not generally entered mainstream diagnostic

convention although the conceptualisation of many psychiatric illnesses as dimensional rather than categorical is increasing and may form an important part of formal diagnostic systems in the future.

1.2.3 Epidemiology

1.2.3.1 Prevalence and incidence

Several epidemiological studies have examined the prevalence and incidence of bipolar disorder around the world. In a synthesis of ten population-based epidemiologic studies using similar methods from the United States, Canada, Puerto Rico, France, West Germany, Italy, Lebanon, Taiwan, Korea, and New Zealand, involving around 38,000 community-dwelling subjects, lifetime rates of bipolar disorder were between 0.3% and 1.5% with the sex ratios nearly equal (Weissman et al., 1996). Other European studies have arrived at similar estimates. In Florence, 1-year and point prevalence estimates for bipolar disorder were 1.7% and 0.6% respectively (Faravelli et al., 1990). In a review of data from fourteen studies (including the aforementioned) from a total of ten countries, the majority of studies reported 12-month estimates of approximately 1% (range 0.5-1.1%), with little evidence of a gender difference (Pini et al., 2005).

The cumulative lifetime incidence of bipolar disorder has been estimated at 1.5-2%, although when the wider range of bipolar spectrum disorders is considered, estimates increase to around 6% (Pini et al., 2005). Within the UK, incidence rates from the recent Aesop study were estimated as 4 per 100,000 per year, ranging from 1.7 in Nottingham to 6.2 in London (Lloyd et al., 2005).

1.2.3.2 Course and outcome

Bipolar disorder has a typical age of onset before the age of thirty. The results of the National Comorbidity Survey from the US suggest that of the four age bands characterised (15 to 24, 25 to 34, 35 to 44, 45 to 54yrs) the 25 to 34yrs group had a significantly greater lifetime odds of a psychiatric illness. For affective disorders specifically, the 12-month odds were greatest in the 15 to 24yrs category and reduced monotonically over time (Kessler et al., 1994). This is consistent with other US studies, such as that by Weissman and colleagues over 10 countries. Here the average age of onset ranged from 17.1 yrs (Edmonton/Alberta) to 27.2 yrs (Puerto Rico) with the first peak between 15 and 19 yrs (Weissman et al., 1996).

The course of the illness is severe and disabling for many patients. In a study conducted through the Stanley Bipolar Foundation Network, 261 patients were screened and completed comprehensive illness and symptom profiling. The average illness duration of those in the sample was 20 years. During the course of their illness, 71% had been hospitalised at least once, 29% had attempted suicide, and 59% had a history of psychosis. Co-morbidity was high with 67% having at least one additional axis-I disorder and 41% had a history of substance abuse. Other important characteristics of the sample were that 52% reported experiencing depressive symptoms first, compared to only 19% for mania/hypomania. Forty six percent reported a worsening of their symptoms over time with only 30% being entirely symptom-free between episodes (Suppes et al., 2001).

As discussed above, although the disorder is characterised by the presence of mania or hypomania, it is depressive symptoms that are predominant. In two prospective natural history studies of weekly symptomatic status, Judd and colleagues performed a long-term follow-up of 146 patients with BP-I over 12.8 years (Judd et al., 2002) and 86 patients with BP-II over 13.4 years (Judd et al., 1998). For patients with BP-I it was found that patients were

symptomatically ill about half of the time (mean, 47.3%) and asymptomatic for the remainder of follow-up (mean, 52.7%). When symptomatic, patients experienced 3 times more depressive symptoms (31.9% of total follow-up weeks) than manic symptoms (9.3% of weeks), and depressive symptoms were 5 times more frequent than cycling/mixed symptoms (5.9% of weeks). For patients with BP-II, again they were symptomatic for about half of the time of the follow-up (mean, 53.9%) however the predominance of depressive symptoms was even greater: patients experienced 39 times more depressive symptoms (50.3% of total follow-up weeks) than hypomanic symptoms (1.3% of weeks), and depressive symptoms were 22 times more frequent than cycling/mixed symptoms (2.3% of weeks).

1.2.4 Summary

From this overview it can clearly be seen that bipolar disorder can for many individuals be seen as a severe and enduring mental illness. One important feature of the illness that emerges is that while the disorder is characterised by (hypo)mania, of the time spent ill, the majority is spent in depression. Therefore it is bipolar depression (BD) that is the focus of this thesis.

1.3 Neuropsychological impairment in bipolar disorder

The following section of this introduction examines the evidence for neuropsychological dysfunction in mood disorders. This will draw from the wider field and will include discussion of impairment in different phases of the illness as well as from depressive disorders in general in order to provide context to the focus on bipolar depression.

1.3.1 General background, methodological issues and clinical correlates relating to neuropsychological impairment in mood disorders

A very brief overview is first presented of some of the broad issues surrounding the assessment of neuropsychological impairments in mood disorders in terms of the factors that have emerged in the literature as being associated with the general profile of impairment. The intention is not to exhaustively review these but to raise those that are relevant to the subsequent literature review and affect the general design of the empirical chapters in this thesis.

1.3.1.1 Unipolar versus bipolar depression

There is a large degree of overlap in the neuropsychological profiles of unipolar and bipolar depression, however in general it appears that there is greater severity of neuropsychological impairment in the latter (Wolfe et al., 1987; Deptula et al., 1991; Borkowska & Rybakowski, 2001). For example, Borkowska and Rybakowski (2001) compared the performance of patients with bipolar or unipolar mood disorders during acute episodes of depression. A significantly greater severity of executive dysfunction – including poorer performance in non-verbal problem solving, response inhibition, verbal fluency and set-shifting – was found in bipolar compared with unipolar depressed patients. Other studies have reported greater executive dysfunction in depressed patients with bipolar disorder (Calev et al., 1986; Martinez-Aran et al., 2004), although it should be noted that others have found only selective impairments in immediate spatial memory and 'hot' (i.e. emotionally-laden) cognition (Roiser et al., 2009) while some found no differences on any measure of attention, memory, executive function, and general intellectual functioning (Mojtabai et al., 2000; Sweeney et al., 2000). The discrepancy is probably due to the confounding effect of other clinical factors. For example, all

subjects in the Mojtabai study were psychotic. Also, it has been reported in one study that while bipolar depressed patients had statistically poorer neuropsychological function than first-episode MDD patients, there was no difference with recurrent patients (Fossati et al., 2004). This highlights the importance of illness variables on performance and raises the question of whether the two disorders can ever be compared satisfactorily. For example, even if the groups could be matched for demographic profile and the severity of current episode, it would be impossible to equate overall illness characteristics such as medication, hospitalizations and previous affective episodes (of depression or mania) (Porter et al., 2007).

1.3.1.2 State versus trait

Neuropsychological impairments persist into euthymia in bipolar disorder (Bearden et al., 2001; Robinson et al., 2006; Bora et al., 2009). There is a general consensus that both executive functions and declarative memory are impaired (van Gorp et al., 1998; Ferrier et al., 1999; Martinez-Aran et al., 2000; El-Badri et al., 2001; Clark et al., 2002; Martinez-Aran et al., 2002; Martinez-Aran et al., 2004; Torrent et al., 2006). However, a recent study suggested that the verbal declarative memory deficits in euthymia may be entirely accounted for by a general executive impairment (Thompson et al., 2009). As is the case with major depression, a number of demographic and clinical characteristics are associated with a more impaired neuropsychological profile in euthymic bipolar patients (Robinson & Ferrier, 2006).

Several studies have compared neuropsychological functioning in euthymic and depressed bipolar patients. It would be logical to assume that the deficits and profile seen in euthymia should also be seen in depression, with the added effect of 'depression' overlaid. For example, Dixon and colleagues reported a similar profile of impairment in executive function between euthymic and depressed bipolar patients, although the error rate on one test of inhibition

(Hayling test; anomalous completion) was greater in the depressed (Dixon et al., 2004). Similarly, Martinez-Aran and colleagues found euthymic and depressed patients to be impaired on most executive and verbal list-learning tasks, but logical memory and visual reproduction was impaired only in the depressed (Martinez-Aran et al., 2004).

However not all studies have found this pattern. Kerr and colleagues reported a similar profile of deficits on the Stroop test in euthymic, mania and depression compared to controls (Kerr et al., 2005). Similarly, in a study by Schneider and colleagues, depressed and euthymic patients did not differ significantly on any verbal or executive domain of the WAIS (Schneider et al., 2008). It is unclear why this pattern emerges in some studies but not others. Subtle differences in clinical or demographic features may play a part but perhaps one important consideration is that it may be a statistical artefact resulting from, in most cases, presenting results in terms of statistical post hoc analysis without clarifying which differences represent a differential deficit (Crawford et al., 2000).

1.3.1.3 Medication effects

Due to the severity of the disorder it is extremely difficult to recruit patients with bipolar disorder who are depressed but drug-free. However, several studies from the USA (Wolfe et al., 1987; Brooks et al., 2006; Taylor Tavares et al., 2007; Roiser et al., 2009) have achieved this (see section 1.3.2). Deficits are seen in these studies although it is difficult to compare these with the overall pattern seen in the area as a whole as there are so few tests in common.

Only one study to date has directly compared medicated and unmedicated bipolar depressed patients (Holmes et al., 2008) and found that there were no statistical differences between

the two, although numerically, accuracy was slightly worse in the unmedicated group, while reaction times were slower in the medicated group. In terms of effect sizes, the unmedicated group performed comparably to those in other studies of medication-free patients, on the tests in common (Roiser et al., 2009). Other studies from India examining neuropsychological performance in unmedicated euthymic patients also reported a similar profile compared to medicated patients (Goswami et al., 2009).

In terms of general impact on neuropsychological functioning, the largest effect to be cognizant of seems to be the psychomotor slowing associated with lithium use (Pachet & Wisniewski, 2003) although even in this area there remains some debate (see Savitz et al., 2005).

1.3.1.4 Clinical and illness modifiers

A number of clinical and demographic factors have been identified as affecting the severity or profile of neuropsychological impairment in mood disorders. Severity of depression, melancholic or psychotic features, co-morbidity, hospitalization, and treatment with ECT have all been associated with a broader and/or more severe impairment (Porter et al., 2007).

In bipolar disorder, neuropsychological functioning has been found to be negatively related to illness features such as the number of episodes suffered, the number of hospital admissions and duration of illness. Episodes of both depression and mania also related negatively to neuropsychological function, although mania was reported to relate more consistently to delayed verbal memory and some measures of executive function, whereas depressive episodes were related less consistently and to a broader range of impairments (Robinson &

Ferrier, 2006). A number of studies have also suggested that the profile of impairment differs in patients with psychotic illness features (Glahn et al., 2006; Levy & Weiss, 2010).

It is unlikely that many of these factors are independent of each other and may simply represent a greater likelihood of impairment the more severe or complex the illness. Also, some of these factors are more easily verified than others, for example, hospitalisation can be easily confirmed whereas precise dates for onset of illness or duration of episode are more difficult to accurately verify. Nevertheless, it is important to be mindful of these features when assessing the profile of impairment between individual studies.

In the next section, the literature on neuropsychological impairment in bipolar depression is reviewed.

1.3.2 Bipolar depression – a review of the neuropsychology literature

Relatively few studies have specifically focused on depression within bipolar disorder. One recent meta-analysis (Kurtz & Gerraty, 2009) which focussed on the neuropsychology of bipolar disorder in euthymia and in symptomatic states (the only one to do this to date) found only 5 papers that met inclusion criteria in the bipolar depression analysis. From these 5 studies, the only tests for which data could be extracted – according to their criteria of requiring similar tests/procedures from at least three – were Trails A (attention/psychomotor speed) and Trails B (executive function), verbal fluency (language) and verbal memory (Rey-AVLT or CVLT). The pooled effect sizes for each of these were d=0.80, 0.64, 0.93, and 1.20 respectively indicate moderate to large effect sizes. A direct comparison with euthymic patients across these measures revealed significantly greater verbal learning and fluency deficits in depressed individuals.

Due to the limited number of studies of sufficient methodological rigour or sharing common tests to permit inclusion in the above meta-analysis, the following section will describe the profile of neuropsychological impairment in bipolar depression in much broader terms. To that end, a database search was conducted in Medline, PubMed, EMBASE and Science Direct to identify suitable papers with a focus on bipolar depression. The terms: 'bipolar disorder' or 'manic-depression' were combined with 'depress*' and 'neuropsycholog*', 'neurocogniti*', 'memory', 'attention', and 'executive'. These were limited to 'English language' and 'Human' studies and were selected from 1980 to present (coinciding with the release of DSM-III). This resulted in the identification of 914 papers. The abstracts of these were assessed to exclude any obviously not focussing on the area of interest. At this stage, review papers were retained to provide additional sources of potential references. Overall, 91 papers were found of interest which were read and assessed for general suitability and reduced to a final selection of 38. The last stage of this process (extraction of mean, s.d. and sample size) resulted in 19

papers being retained with sufficient data of use; of the remainder, five were general review papers and fourteen had issues with the data or methodology e.g. no control group, only mild depressive symptoms, pooled analysis with other diagnoses or insufficient data to calculate an effect size.

The following section summarises the results of these studies. General information is also summarised alongside the neuropsychological data; specifically other groups included in the studies, demographic details, illness severity and diagnostic procedures, current hospitalisation and medication use, and inclusion/exclusion criteria. Primarily the focus is on the major broad domains of neuropsychological function; attention and executive function, immediate/ short-term memory, verbal memory, visuo-spatial memory, (psycho)motor speed. Results will be described on a study by study basis but where multiple studies have used the same or similar tasks, a meta-analytic approach will be taken to pool the effect sizes and give an overall estimate with confidence intervals [1]. Where possible this will be done on an individual test-by-test basis and by a broader 'process' approach e.g. immediate verbal recall, delayed verbal recognition etc. The layout of the data is in a form where negative effect sizes always indicate lower/worse performance in patients compared to controls and similarly, significance of the pooled effects. Effect sizes included in the pooled analyses are indicated in the tables with an asterisk.

[1] For the pooled estimate of effect size within a meta-analysis, both random and fixed-effects methods can be produced. As outlined in the Cochrane Collaboration methodology there is considerable debate as to the most appropriate to use. In the current data, in the majority of cases the fixed and random effects produced identical pooled effect estimates, however, when there were differences the fixed effect model tended to produce a marginally more conservative estimate. Therefore the fixed effect method is used throughout.

Table 1-1 Overview of the 19 studies identified.

ID	BD (n)	CON (n)	Other groups (n)	BD (sex/age)	Controls (sex/age)	Other (sex/age)	Diagnosis	Depression severity	Hospital	Medication	Matching	Inclusion/exclusion	TESTS	Summary/notes
1	25	34	manic= 37, mixed= 24	7m,18f age=35.2, s.d.=11	3m,31f age=34.1, s.d.=13.9	manic: 11m,26f age=35.7, s.d.=10.43; mixed: 3m, 21f age=36.0, s.d.=8.7	DSM-IV	Not reported	Inpatient	8 drug-free; all others medicated	Age, gender, education	loss of consciousness, neurological disease	CVLT, Verbal fluency, Trails, Grooved pegboard	Statistically, all groups were impaired vs controls.
4	8	27	n/a	sex cannot be calculated. Age=37.1, s.d.=7.8	16m,11f age=34.0, s.d.=11.7	n/a	Not stated	HAMD=32.8, s.d.=7.4	Inpatient	Drug free (2 weeks)	Not reported	seizure disorder; progressive neurological/ systemic disorder; unstable medical illness; administration of concomitant medication (e.g. benzodiazepines, antidepressants, mood stabilizers, stimulants, steroids); hormone replacement therapy; electroconvulsive or light therapy; administration of any investigational drug within 30 days; abuse of any alcohol (2 weeks); abuse of any drugs (<6 months); incapable of understanding/ consenting. Controls: 1st degree rel.	Connor's CPT	Done in the context of fMRI. Paper reports that patients were non-statistically worse than controls on all measures.
5	24	24	n/a	15m,9f age=40.0, s.d.=9.5	10m,14f age=34.0, s.d.=9.1	n/a	DSM-IV SCID	HAMD=33.7, s.d.=7.1	Not stated	Mood stabilisers (no antipsychotics or antidepressants)	Sex, handedness	Patients: HAMD<18. substance abuse (1 month)/ dependence (6-month); rapid cycling (1 year); concomitant Axis I disorder; current antipsychotic medication; current antidepressant/ stimulant use; and use of benzodiazepines (6 h of testing).	d2-Concentration, Cued Target Detection, Stroop, Finger tapping test, Grooved pegboard, Simple reaction time, Choice reaction time	Statistically, bipolar patients' performance was significantly slower than controls on a single psychomotor measure, involving a cognitive speed component. On tasks that measure pure motor speed, the performance of patients did not differ.
7	8	14	MDD= 20	Sex for BD group not reported. Age=31.1, s.d.=4.5	Age=34.5, s.d.=5.5	MDD: age=35.5, s.d.=6.4	DSM-III	HAMD=23.4, s.d.=10.8	Inpatient	"2/3" on medication (started 1-2wks prior)	Age, sex, education	Brain damage, epilepsy, ECT, drug/alcohol abuse, "penal incarceration"	Verbal and non-verbal memory (recall and recognition)	Also divided overall depressed group into DST suppressors and non-suppressors. Concluded that depressed patients show asymmetry of recall, with non-verbal worse than verbal. Worse in bipolar patients and non-suppressors.

ID	BD (n)	CON (n)	Other groups (n)	BD (sex/age)	Controls (sex/age)	Other (sex/age)	Diagnosis	Depression severity	Hospital	Medication	Matching	Inclusion/exclusion	TESTS	Summary/notes
8	15	30	manic=15, remitted BD=15	6m,9f age=33.9 ,s.d.=8.2	17m,13f age=35.2 , s.d.=9.8	Manic: 7m,8f age=34.3 , s.d.=11.6 ; remitted: 8m,7f age=35.7 , s.d.=9.3	DSM-IV	BDI=29.7, s.d.=10.1	Outpatients	All on medication	Age, sex, ethnicity, socio-economic status, handedness (Y.ed. differed but only in remitted group)	personal history of drug/alcohol abuse/ neurological disorder (all subjects); personal/ family history of psychiatric illness (controls); symptoms too severe; level of cooperation too low to provide valid data (patients). English first language.	Phonological & Semantic Verbal Fluency, Hayling Test, Stroop, Cognitive Estimates, WAIS-R Vocabulary, Rate of articulation	Statistically there were significant differences between depressed and controls on Hayling and Stroop. Greatest impairment in manic, esp. thought disordered.
10	18	88	feMDD=23, recurrMDD=28	7m,11f age=42.3, s.d.=10.3	38m,50f age=43.7 , s.d.=14.8	feMDD: 7m,16f age=40.1 , s.d.=14.8 ; recurrMDD: 4m,24f age=43.1 , s.d.=14.8	MINI and DSM-IV checklist	MADRS=26 .8, s.d.=6.3	Inpatient	All on medication	Age, education	Concurrent Axis II diagnoses, substance abuse, brain diseases likely to affect cognition, e.g. dementia, Parkinson's disease, stroke, head injury, ECT <12mths. All recruited within 1 week of admission.	Memory (verbal), modified WCST, Digit span	FE did not show Memory (verbal) impairment vs controls. UR and BP patients exhibited Memory (verbal) deficits with impaired free recall and normal cued recall and recognition. Deficits of the UR and BP present in the first free recall trial. Depressed patients improved memory performance across the three trials of the task at the same rate as controls.
11	BDnp =15, BDp= 11	32	schizoaffective=15, schizophrenia=15	BDnp: 4m,11f age=37.3, s.d.=10.6; BDp: 3m,8f age=35.36, s.d.=10.4	13m,19f age=39.0 , s.d.=10.5	schizoaffective: 6m,9f age=36.4 , s.d.=10; schizophrenia: 8m,7f age=37.5 , s.d.=6.8	DSM-IV	BDnp: HAMD=10. 0, s.d.=6.7; BDp: HAMD=14. 3, s.d.=3.0	Outpatients	All on medication	Age, sex and race	Nno concomitant Axis I disorder; history of an illness/ event that might affect cognitive function (e.g., epilepsy, migraine, significant head trauma). Controls: no first degree relatives.	Digit span, Spatial DRT	BDp and BPnp and those schizophrenia/schizoaffective disorder impaired on backward digit span. Only patients with a lifetime history of psychotic features, regardless of diagnosis, were impaired on spatial DRT. (n.b. IQ and education different in SCZ groups but not BD and vs controls)
14	UB*P*= 32, MB*P*= 33	52	n/a	UBP: 8m,24f age=35.3, s.d.=8.7; MBP: 15m,18f age=41.1, s.d.=10.9	27m,25f age=37.0 , s.d.=10.1	n/a	DSM-IV	UBP: MADRS=25 .7, s.d.=8.3; MBP: MADRS=32 .1, s.d.=5.8	Outpatients	Medicated: stable on lithium (n=23) or valproate (n=10) for >4 weeks; all other medications were discontinued >4 weeks. Unmedicated: 3-6 weeks.	Age, and IQ	Not explicitly stated. No statistical differences in IQ and age but greater proportion of females in UBP vs MBP and controls.	SWM, PRM , RVIP	Statistical differences reported in RVIP omission errors (MBD > HC) and Affective Shift task (Happy shift omissions: MBD >UBD, HC; All shift latency: MBD >UBD, HC) although data cannot be extracted for analysis. Study compares medicated and unmedicated depressed bipolar patients. Matching (MADRS: MBD > UBD > HC Length of illness: MBD=UBD. Females: UBD > MBD, HC)

ID	BD (n)	CON (n)	Other groups (n)	BD (sex/age)	Controls (sex/age)	Other (sex/age)	Diagnosis	Depression severity	Hospital	Medication	Matching	Inclusion/exclusion	TESTS	Summary/notes
17	13	18	euthymic=15, manic=14, MDD=17	6m,7f age=49.7, s.d.=11.5	8m,10f age=41.8, s.d.=13.1	Euthymic: 6m,9f age=47.7, s.d.=9.1; Manic: 8m,6f age=43.1, s.d.=15.9; MDD: 8m,9f age=42.8,s.d.=13.2	DSM-IV	BDI=25.7, s.d.=11.8	In and Outpatients	All on medication	age, gender, and IQ	substance abuse, medical conditions that impair neurological function, colour-blindness, MMSE<24	Card Stroop, Emotional Stroop	All patients were significantly more impaired than normal healthy controls on all conditions. No differences between the groups.
19	15	30	Schizophrenia; thought disorder (n=15), negative symptoms (n=15), manic=15	6m,9f age=33.9, s.d.=8.2	17m,13f age=35.2, s.d.=9.8	TD=11m,4f age=33.4, s.d.=11.0, NS=10m,5f age=34.1, s.d.=8.1, manic=7m,8f age=34.3, s.d.=11.6	DSM-IV	BDI=29.7, s.d.=10.1	??	All on medication	age, gender, ethnicity, handedness, socio-economic status, years of education	drug/alcohol abuse or neurological disorder, English as first language.	Phonological and semantic VF, HSCT Stroop, Cognitive estimations test ,WAIS-R vocabulary	Aim was to see if symptom pattern ('excess' {disorganisation/mania} or 'deficiency' {negative symptoms/depression]} would be more related to executive ability than the underlying disorder. No differences between groups with 'excess' symptoms (schizophrenia patients with thought disorder and bipolar patients with mania), or between groups with 'deficiency' symptoms (schizophrenia patients with negative symptoms and bipolar patients with depression). In contrast, differences were noted between groups with the same diagnosis: Schizophrenia patients with disorganisation were less accurate in semantic verbal fluency than those with negative symptoms; and bipolar patients with mania tended to be faster, but less accurate, in sentence completion than those with depression.
25	30	30	manic/hypomanic=34; euthymic=44	15m,15f age=43.4, s.d.=10.7	8m,22f age=38.9, s.d.=12.4	manic/hypomanic=17m,17f, age=42.4, s.d.=11.9; euthymic=18m,26f, age=39.6, s.d.=9.5	DSM-IV bipolar I or II depressed	HAMD=19.7, s.d.=3.2	??	All on medication	Age, sex, level of education	HAMD≥17, physical or neurological illness, a history of head injury, neurodegenerative disorder, substance abuse/ dependence in the last year, mental retardation, ECT in the last year. Controls-also 1st deg relatives	WCST, Stroop, Digit span, Trails, FAS Category fluency, (animals) CVLT, WMS-R logical memory and Nonverbal (visual reproduction)	Depressed group statistically worse on every test and measure

ID	BD (n)	CON (n)	Other groups (n)	BD (sex/age)	Controls (sex/age)	Other (sex/age)	Diagnosis	Depression severity	Hospital	Medication	Matching	Inclusion/exclusion	TESTS	Summary/notes
27	24	62	MDD=27, dysthymic=14, schizoaf*f*=15	14m,10f age= 49.58, s.d.=7.72	18m,44f age=52.3, s.d.=7.15	MDD: age= 49.81, s.d.=8.75; dysthymic: age=46.29, s.d.=8.26; SA: age=47.27,s.d.=9.91	DSM-IV	HAMD done but not reported	Inpatient	36.6% drug free	Age, education	acute intoxication; delirium; heavy sensory or motor handicap; other than German native language; ECT 12mths; current/ recent substance abuse; history of CNS disease	REY-AVLT (modified) , Trails, Verbal fluency, WMS	Study examines performance at admission and discharge of several groups. Bipolar group had improved significantly in the first trial of Memory (verbal), verbal fluency and depression score.
28	24	30	MDD=46	9m,15f age=35.4, s.d.=??	9m,21f, age=34.0, s.d.=??	MDD: 16m,30f age=37.0, s.d.=??	Perris's criteria	HAMD=20.3, s.d.=9.5	Inpatient	??	Age, sex, education	??	Simple/ motor reaction time, choice RT, memory scanning, Stroop, counting task, tapping speed	Concluded that generally the unipolar group was more impaired than the bipolar group
30	49	55	n/a	13m,36f age=33.6, s.d.=8.9	19m, 36f age=34.9, s.d.=8.1	n/a	DSM-IV (BD-I: 11, BD-II: 38)	MADRS=24.8, s.d.=10.2	Outpatients	Drug-free (3 weeks), 8 for fluoxetine	Age, sex, IQ	major medical illness/ neurological disorder; history of head injury; substance abuse (<6 mths); substance dependence (<5 yrs); electrolyte disturbance; anaemia; positive urine drug screen. Additional exclusion criteria for the BD: age at onset for the first mood episode of >40 years or psychotic manifestations	WASI, CANTAB SSP, SWM, IDED, SREC, PREC, DMTS, RVIP, Gambling, PRL, affective Go/no go	Greatest impairment in 'hot' cognition. In other measures, no significant difference except BD worse on spatial span.
31	24	26	n/a	Gender unknown; age=43.7, s.d.=2.3	Gender unknown; age=39.3, s.d.=2.5	n/a	DSM-IV BD-I	HAMD17 =25.2, s.d.=1.5	15 in-, 9 outpatients	On medication	Age and IQ	ECT (<6mths), history of learning disability/ neurological illness/ unstable medical illness affecting cognition, current/ past alcohol /drug dependence	CANTAB SMTS DMTS PREC SREC IDED TOL Affective go/no-go, Decision-making task	Significant impairments in BD on all accuracy measures except SMTS and TOL. Also examined proportion <5th %ile.
34	32	28	34 euthymic BD	4m,28f; age=45.2, s.d.=10.5	16m, 12f age=35.1, s.d.=14.1	euthymic BD: 14m,20f age=42.3, s.d.=13.1	DSM-IV	HAMD =15.5, s.d.=5.5	Outpatients	n=2 drug-free	None stated (age, IQ, sex differed)	ECT/alcohol or drug abuse <6months, neurological disease, hypomania, 'literate'	WAIS-III (verbal, executive and factors)	In terms of statistical significance, patients impaired on all factors and most individual sub-tests (Euthymic/depressed groups defined as > / < 8 HAMD) n.b. effect sizes reported in paper smaller than here as age and IQ added as covariates

ID	BD (n)	CON (n)	Other groups (n)	BD (sex/age)	Controls (sex/age)	Other (sex/age)	Diagnosis	Depression severity	Hospital	Medication	Matching	Inclusion/exclusion	TESTS	Summary/notes
36	21	51	14 mixed/manic BD, 58 MDD	12m:9f, age=31.9, s.d.=1.4	12m:39f, age=36.3, s.d.=9.7	mixed/manic: 6m, 8f; age=36.1, s.d.=11.0 MDD:19m,39f; age=32.3, s.d.=9.1	DSM-IV BD	HAMD17=17.3, s.d.=5.46	Inpatient	On medication	Age and IQ	ECT, systemic neurologic disease, recent substance abuse	CANTAB Big Circle/Little Circle, FSRT, SOC, SWM, IDED, SREC, PREC, DMTS, MTS, PAL, SSP	In terms of statistical significance, no differences reported between BD and controls on any measure. (n.b. SWM Between errors only reported as figure for levels 4 to 8)
37	17	25	22MDD	5m:12f; age=32.6, s.d.=11.1	7m:18f; age=34.8, s.d.=8.8	5m:17f; age=38.6, s.d.=8.1	DSM-IV BD-II	MADRS=24.1, s.d.=9.73	Outpatients	Drug-free (3 weeks), 8 for fluoxetine	Age and IQ	alcohol/substance abuse (<1 year), lifetime dependence, neurological disease, head injury, current pregnancy, a full scale IQ <85	CANTAB PRM, SRM, SSP, S/DMTS, SWM, IDED, CGT, IST	General summary that bipolar II depression not associated with neuropsychological impairment, unlike UP (but n.b. Pairwise analysis only done if the anova main effect was significant)
38	12	20	20UP; 10HD	12m; age=48.3, s.d.=14.1	13m:7f; age=49.8, s.d.=10	UP: 19m, 1f; HD: 6m, 4f	DSM-III BD-I	HAMD17=23.8, s.d.=7.6	Inpatient	Drug-free (1 week)	Age and education	Head injury, ECT, alcohol/drug abuse, neurologic disorder, learning difficulties	Rey-AVLT, Verbal fluency	BD worse that controls on verbal learning, but not delayed recall (nb. Not standard administration) and verbal fluency. BP worse than UP patient on all measures.

Key to references: 1. Basso, M. R., et al. 2002; 4. Brooks, J. O., et al. 2006; 5. Burdick, K. E., et al. 2009; 7. Deptula, D., et al. 1991; 8. Dixon, T., et al. 2004; 10. Fossati, P., et al. 2004; 11. Glahn, D. C., et al 2006; 14. Holmes, M. K., et al. 2008; 17. Kerr, N., et al. 2005; 19. Kravariti, E., et al. 2005; 25. Martinez-Aran, A., et al. 2004; 27. Neu, P., et al. 2001; 28. Popescu, C., et al. 1991; 30. Roiser, J. P., et al. 2009; 31. Rubinsztein, J. S., et al. 2006; 34. Schneider, J. J., et al. 2008; 36. Sweeney, J. A., et al. 2000; 37. Taylor Tavares, J. V., et al 2007; 38. Wolfe, J., et al. 1987

Key to abbreviations: CGT: CANTAB Cambridge Gambling Test; CPT: Continuous Performance Test; CVLT: California Verbal Learning Test; DMTS: CANTAB Delayed-Match-To-Sample; DRT: Delayed Response Test; FAS; 'FAS' Verbal Fluency; FE: first-episode; FSRT: CANTAB Five Stage Reaction Time; HC: Healthy Controls; HSCT: Hayling Sentence Completion Test; IDED: CANTAB Intra-dimensional/Extra-dimensional Shift; IST: Information Sampling Test; MBP: Medicated Bipolar Patients; MDD: Major Depressive Disorder; PAL: CANTAB Paired Associative Learning; PRL: Probabilistic Reversal Learning; PRM/PREC: CANTAB Pattern

Tabulated above are the 19 papers identified, including sample sizes, basic demographics, diagnostic and illness characteristics/ severity measures, inclusion and exclusion criteria, tests employed and a general summary of the outcome of the study. The average sample size in the studies is n=24 (range, n=8 to 65) for the bipolar depressed group and n=36 (range, n=14 to 88) for controls. This slight disparity is due to some studies having additional experimental groups of interest e.g. MDD, manic/euthymic bipolar groups etc. Most groups are matched (or are at least non-statistically different) in age, sex and some measure of IQ or general educational attainment. Due to the severe nature of the illness, most patients were receiving medication at the time or testing although some studies have been in individuals free of psychotropic medications (Wolfe et al., 1987; Brooks et al., 2006; Taylor Tavares et al., 2007; Roiser et al., 2009) or have been designed to directly test the differences between medicated and unmedicated patients (Holmes et al., 2008).

One of the most striking issues with the studies listed is the heterogeneity in selection of neuropsychological tasks. As was suggested in the Kurtz & Gerray (2009) meta-analysis, there are very few studies that have used the same or similar measures within any one theoretical domain. For this reason, a less stringent criteria will be used for assessing and pooling data. If two or more studies have used the same or similar measures, these will be examined together. For clarity, this will be explained on a test by test basis in the subsequent sections.

1.3.3 Executive and attentional impairments

Although it is noted that some theoretical models of human memory include the control of attention as an executive function (Baddeley & Hitch, 1974) the two will be considered separately in this section[2].

1.3.3.1 Attention

Due to the nature of the task, the majority of studies have utilised computerised tests to assess aspects of sustained attention. In the Conner's CPT, subjects view letters presented in the centre of a screen, one at a time, for varied presentation rates and intervals. Subjects are instructed to respond as fast and as accurately as possible every time a letter appears on the screen, except for the letter 'X'. The entire test lasts approximately 15 minutes. The CANTAB RVIP (rapid visual information processing) task is analogous to this, with subjects viewing series of numbers on the screen and are required to respond each time they see one of three target sequences (3-5-7, 2-4-6, or 4-6-8). The overall test lasts 4 minutes. The cued target detection paradigm is a "Posner-style" pre-cuing task in which subjects are required to detect the presence of an 'x' inside of one of two boxes whilst fixating on a central '+'. Just prior to the target appearing in one of the boxes, subjects were explicitly told to "watch for a cue which would correspond with the side of the screen in which the target would appear 80% of the time". Subjects were then instructed to respond to the target as quickly as possible by pressing a key on the side of the keyboard corresponding with the side of the screen on which the target appeared. Cues were of varying types, including valid, invalid and neutral. Finally, the d2 task, the only pen-and-paper test requires subjects to cross out any letter 'd' that has two flankers (dashes) in any combination.

[2] For the purposes of these initial chapters, the attribution of tasks into neuropsychological domains is done according to common classification in the neuropsychological literature (for example, Lezak et al., 2004; Strauss et al., 2006).

Table 1-2. Summary of studies examining attentional tasks in bipolar depression

Paper	Study		BD			Con			TEST	Outcome measure
		Cohen's d	n	mean	s.d.	n	mean	s.d.		
4	Brooks, J. O., et al. 2006	-0.084	8	56.4	14.7	27	55.1	15.6	Connor's CPT	Beta
		-0.752	8	55	18.9	27	45.6	10.1	Connor's CPT	Commission errors
		-0.610	8	54.5	15.5	27	45.8	13.9	Connor's CPT	d'
		-0.084	8	56.4	14.7	27	55.1	15.6	Connor's CPT	Hit rate RT
		-0.120	8	47.8	21.1	27	49.6	12.8	Connor's CPT	Hit RT (block change)
		-0.258	8	61.7	17	27	57.3	17.1	Connor's CPT	Omission errors
5	Burdick, K. E., et al. 2009	0.346	24	44.1	36.4	24	32.6	29.7	Cued Target Detection	Benefit 100 ms
		0.525	24	47.7	59.8	24	23.4	26.7	Cued Target Detection	Benefit 400 ms
		0.435	24	8.9	40.3	24	-9.0	41.9	Cued Target Detection	Benefit 800 ms
		-0.265	24	-75.4	191.7	24	-38.2	51.6	Cued Target Detection	Cost 100 ms
		-0.408	24	-88.2	152.3	24	-42.3	46.5	Cued Target Detection	Cost 400 ms
		-0.427	24	-23.8	75.8	24	2.4	42.3	Cued Target Detection	Cost 800 ms
		-0.958	24	171.46	49.5	24	214.17	39.1	d2-Concentration	Total
30	Roiser, J. P., et al. 2009	-0.222	49	0.92	0.05	55	0.93	0.04	RVIP (CANTAB)	A'
		-0.180	49	0.96	0.05	55	0.97	0.06	RVIP (CANTAB)	B''
14	Holmes, M. K., et al. 2008	-0.154	32	1.0	1.3	52	0.8	1.3	RVIP (CANTAB)	Commission errors
		0.082	33	0.7	1.1	52	0.8	1.3	RVIP (CANTAB)	Commission errors
30	Roiser, J. P., et al. 2009	0.216*	49	462	89	55	493	179	RVIP (CANTAB)	Latency
14	Holmes, M. K., et al. 2008	0.173*	32	454.3	66.3	52	469.4	97.9	RVIP (CANTAB)	Latency
		-0.416*	33	509	90.4	52	469.4	97.9	RVIP (CANTAB)	Latency
		-0.469	32	8.9	5.4	52	6.7	4.2	RVIP (CANTAB)	Omission errors
		-0.991	33	11.1	4.8	52	6.7	4.2	RVIP (CANTAB)	Omission errors

Within this domain there is a great deal of heterogeneity in the number of outcome measures that have been used, producing highly variable effect size estimates. The only measure used in two separate studies is the latency measure from the RVIP (Holmes et al., 2008; Roiser et al., 2009). As the Holmes et al study involved the comparison of medicated and unmedicated patients, two separate comparisons are included from this study, although the same control group is included in both. The pooled estimate of effect size for RVIP latency was d= -0.011 (95%CI=0.231 to -0.254; χ^2=0.008, *p*=0.928; individual effect size plots are presented in appendix 9.1). However, it is of note that other single measures produced much greater effect sizes such as commission errors (d= -0.752) in Connor's CPT (Brooks et al., 2006) and RVIP omission errors (d= -0.991) in medicated patients (Holmes et al., 2008). It is also notable that the largest single effect size in this domain was for the d2 task (d= -0.958), the only pen-and-paper task employed, with a greater psychomotor/fine dexterity component than other measures.

1.3.3.2 Executive functioning

Due to the complexity of defining executive functions (Lezak et al., 2004; Strauss et al., 2006), this neuropsychological domain contains the greatest number of tests tapping many different theoretical executive processes. Those used here can fit generally into the domains of working memory monitoring; set-shifting/ rule formation and reversal; planning, reasoning and strategy; inhibition; and fluency. Although tasks can often fit into multiple domains, for the purpose of obtaining a pooled estimate of effect size, each outcome measure is only used once.

Tests within the *Working memory monitoring* domain include: the digit span (reverse), where participants must immediately recall a string of numbers in the reverse order from how they were presented; the Sternberg memory scanning test, where participants are required to rapidly scan strings of letters and indicate the presence or absence of a target; and the CANTAB spatial working memory (SWM) test, in which participants must carry out a self-ordered search for target 'tokens' hidden in an array of boxes.

Tests within the *Set shifting/ rule formation and reversal* domain include: the Wisconsin Card Sorting Test (WCST) and the analogous CANTAB intra-dimensional/extra-dimensional (IDED) set-shifting task, involve the formation and subsequent shifting of rules that change through the viewing or sorting of visual stimuli. Similarly, Big Circle/Little Circle (CANTAB) tests reaction time and the ability to follow and then reverse rule-based responses; Trails B, which requires participants to switch between letters and numbers, joining one to the other.

Within the domain of *Planning, reasoning and strategy* the tests employed were: CANTAB Stockings of Cambridge or Tower of London (SOC/TOL), where participants must plan and execute a number of moves to rearrange a series of coloured disks into a given arrangement; the Cognitive Estimation Test, which requires participants to estimate the answers to questions using deductive reasoning; and the strategy score from the CANTAB SWM, which assesses the efficiency of the search strategy adopted.

Tests of *Inhibition* included: the Stroop test, in which participants must inhibit the pre-potent response of reading the printed names of colours and instead must say the colour of the ink that the word is printed in; the Hayling Sentence Completion Test (HSCT), which requires participants to finish incomplete sentences using contextually relevant (part A: straightforward completion condition) or contextually irrelevant (part B: anomalous

completion condition) words i.e. inhibit the obvious response; the CANTAB go/no-go task, and perseverations from the WCST .

Finally, the verbal *fluency* tests that were included were either phonological (i.e. words beginning with a given letter) or by category (e.g. animals).

Table 1-3. Summary of studies examining attentional tasks in bipolar depression

Paper	Study		BD			Con			TEST	Outcome measure
		Cohen's d	n	mean	s.d.	n	mean	s.d.		
36	Sweeney, J. A., et al. 2000	-0.375*	21	809	152	51	758	129	Big Circle/Little Circle (CANTAB)	Response latency (msec)
8	Dixon, T., et al. 2004	-0.437*	15	8.9	5.1	30	6.9	4.3	Cognitive Estimates Test	Error Score
10	Fossati, P., et al. 2004	-0.979*	17	4.2	0.97	49	5.4	1.3	Digit span	Reverse
25	Martinez-Aran, A., et al. 2004	-0.909*	30	3.8	1.1	30	4.8	1.1	Digit span	Reverse
11	Glahn, D. C., et al 2006	-0.766*	15	6.03	1.6	32	7.22	1.7	Digit span	Reverse
		-1.334*	11	5.12	1.4	32	7.22	1.7	Digit span	Reverse
8	Dixon, T., et al. 2004	-1.285*	15	20.5	11.4	30	10.5	5.2	Hayling Test	Error score (anomalous completion)
		-0.876	15	0.3	0.6	30	0	0	Hayling Test	Error score (straightforward completion)
		-1.269	15	28.8	20.3	30	13.6	3.7	Hayling Test	Response Initiation Latency
		-0.463	15	31.2	35.7	30	20.4	13.9	Hayling Test	Response Suppression Latency
30	Roiser, J. P., et al. 2009	-0.253*	49	11.1	10.3	55	8.6	9.5	IDED (CANTAB)	Errors ED stage
31	Rubinsztein, J. S., et al. 2006	-0.418*	24	16.5	11.8	26	12.3	8.16	IDED (CANTAB)	Errors ED stage
36	Sweeney, J. A., et al. 2000	-0.152*	21	5.38	7.19	51	4.4	6.13	IDED (CANTAB)	Errors ED stage (adjusted)
31	Rubinsztein, J. S., et al. 2006	0.232	24	7.4	4.9	26	8.9	7.69	IDED (CANTAB)	Errors pre-ED stage
36	Sweeney, J. A., et al. 2000	-0.299	21	9.24	5.84	51	7.44	6.09	IDED (CANTAB)	Errors pre-ED stage
31	Rubinsztein, J. S., et al. 2006	-0.985	24	8.04	1.0	26	8.8	0.51	IDED (CANTAB)	Stages completed
37	Taylor Tavares, J. V., et al 2007	0.075	17	8.87	2.14	25	8.76	0.7	IDED (CANTAB)	Stages completed
36	Sweeney, J. A., et al. 2000	-0.549	21	8.43	1.72	51	9.0	0.57	IDED (CANTAB)	Stages completed
30	Roiser, J. P., et al. 2009	-0.227	49	18.8	11.2	55	16.2	11.7	IDED (CANTAB)	Total errors
37	Taylor Tavares, J. V., et al 2007	-0.251*	17	16.5	9.90	25	13.9	10.65	IDED (CANTAB)	Total errors
30	Roiser, J. P., et al. 2009	0.060	49	5.5	4.9	55	5.8	5.1	IDED (CANTAB)	Total reversal errors

Paper	Study		BD			Con			TEST	Outcome measure
		Cohen's d	n	mean	s.d.	n	mean	s.d.		
28	Popescu, C., et al. 1991	-0.932*	24	30.1	10.89	30	22.58	4.75	Memory scanning	digits 1
		-1.463	24	39.99	13.38	30	26.22	4.11	Memory scanning	digits 2
		-1.068	24	34.09	11.54	30	25	4.92	Memory scanning	letters 1
		-1.375	24	46.76	15.03	30	31.86	5.61	Memory scanning	letters 2
		-1.166	24	65.34	22.6	30	46.2	8.84	Memory scanning	letters 3
		-1.012	24	79.06	27.37	30	58.22	12.9	Memory scanning	letters 4
28	Popescu, C., et al. 1991	-1.068	24	86.06	21.06	30	68.79	10.83	Stroop Test	colour naming
17	Kerr, N., et al. 2005	-1.235	13	58.09	9.85	18	72.88	13.28	Stroop Test	colour naming
17	Kerr, N., et al. 2005	-1.278*	13	39.72	7.1	18	52.64	11.78	Stroop Test	colour-word score
8	Dixon, T., et al. 2004	-1.192*	15	70.8	31.6	30	97.5	16.2	Stroop Test	colour-word score
5	Burdick, K. E., et al. 2009	-0.606*	24	0.63	6.5	24	5.25	8.6	Stroop Test	interference
25	Martinez-Aran, A., et al. 2004	-1.058*	30	-2.3	6.6	30	4.9	7	Stroop Test	interference
28	Popescu, C., et al. 1991	-1.302*	24	146.56	38.08	30	107.52	21.52	Stroop Test	interference
		0.165	24	49.29	8.41	30	51.05	12.2	Stroop Test	word reading
17	Kerr, N., et al. 2005	-0.722	13	85.63	17.17	18	95.47	10.42	Stroop Test	word reading

Paper	Study		BD			Con			TEST	Outcome measure
		Cohen's d	n	mean	s.d.	n	mean	s.d.		
30	Roiser, J. P., et al. 2009	0.242	49	0.73	1.9	55	1.3	2.7	SWM (CANTAB)	Between error (4-box)
		-0.242	49	6.8	8.1	55	5	6.8	SWM (CANTAB)	Between error (6-box)
		-0.046	49	16	14	55	15.4	12	SWM (CANTAB)	Between error (8-box)
14	Holmes, M. K., et al. 2008	-0.456*	32	28.1	22.4	52	18.9	18.7	SWM (CANTAB)	Between error (total)
		-0.333*	33	24.9	16.8	52	18.9	18.7	SWM (CANTAB)	Between error (total)
37	Taylor Tavares, J. V., et al 2007	-0.501*	17	25.6	19.79	25	15.9	19.05	SWM (CANTAB)	Between error (total)
14	Holmes, M. K., et al. 2008	-0.409*	32	33.5	7	52	31	5.5	SWM (CANTAB)	Strategy
		-0.172*	33	32	6.3	52	31	5.5	SWM (CANTAB)	Strategy
30	Roiser, J. P., et al. 2009	-0.162*	49	32.6	6.7	55	31.6	5.7	SWM (CANTAB)	Strategy
36	Sweeney, J. A., et al. 2000	-0.600*	21	36.1	4.41	51	33.2	4.99	SWM (CANTAB)	Strategy
37	Taylor Tavares, J. V., et al 2007	-0.633*	17	33.2	5.77	25	29.5	5.9	SWM (CANTAB)	Strategy
30	Roiser, J. P., et al. 2009	0.173	49	0.11	0.75	55	0.23	0.64	SWM (CANTAB)	Within error (4-box)
		-0.169	49	0.73	2.4	55	0.42	1.1	SWM (CANTAB)	Within error (6-box)
		-0.064	49	1.8	3	55	1.6	3.2	SWM (CANTAB)	Within error (8-box)
14	Holmes, M. K., et al. 2008	-0.234	32	3	5.5	52	2	3.3	SWM (CANTAB)	Within errors (total)
		0.030	33	1.9	3.4	52	2	3.3	SWM (CANTAB)	Within errors (total)

Paper	Study		BD			Con			TEST	Outcome measure
		Cohen's	n	mean	s.d.	n	mean	s.d.		
36	Sweeney, J. A., et al. 2000	-0.535*	21	7.19	1.63	51	8.16	1.88	TOL/SOC (CANTAB)	Minimum moves
31	Rubinsztein, J. S., et al. 2006	-0.285*	24	0.58	0.1	26	0.64	0.25	TOL/SOC (CANTAB)	Proportion correct
1	Basso, M. R., et al. 2002	-0.808*	25	90.25	48.38	34	60.73	24.61	Trails	B
25	Martinez-Aran, A., et al. 2004	-0.863*	30	151.2	113.9	30	77.7	39.1	Trails	B
		-0.918*	30	16.8	5.1	30	21.3	4.7	Verbal fluency (animals;	Correct
27	Neu, P., et al. 2001	-0.965*	24	25.04	6.36	62	32.81	8.6	Verbal fluency (animals;	Correct
1	Basso, M. R., et al. 2002	-0.773*	25	36.12	11.47	34	44	9.16	Verbal fluency (FAS)	Correct
25	Martinez-Aran, A., et al. 2004	-1.230*	30	25.3	12.6	30	39.9	11.1	Verbal fluency (FAS)	Correct
38	Wolfe, J., et al. 1987	-1.114*	12	-	-	20	-	-	Verbal fluency (FAS)	Correct
8	Dixon, T., et al. 2004	-0.640*	15	87.5	10.6	30	92.7	6.6	Verbal Fluency (Phonological)	% correct
		-0.563	15	32.3	18.5	30	40.3	11.6	Verbal Fluency (Phonological)	Total responses across trials
		-0.416*	15	95.4	5.9	30	97.2	3.3	Verbal Fluency (semantic)	% correct
		-0.180	15	39.3	1.6	30	39.6	1.7	Verbal Fluency (semantic)	Total responses across trials
34	Schneider, J. J., et al. 2008	-1.139	32	5.4	2.1	28	8.9	3.9	WAIS-III (executive)	complete figures
		-1.317	32	6.2	2.4	28	9.3	2.3	WAIS-III (executive)	cubes
		-1.145	32	7.9	2.5	28	11.1	3.1	WAIS-III (executive)	matrix reasoning
		-0.787	32	6.6	2.2	28	9	3.8	WAIS-III (executive)	picture arrangement
25	Martinez-Aran, A., et al. 2004	-0.514*	30	4.6	1.7	30	5.4	1.4	WCST	Categories
		-1.084*	30	18.9	10.4	30	9.2	7.2	WCST	Perseverative errors
10	Fossati, P., et al. 2004	-0.276*	17	4.6	1.72	49	5	1.35	WCST (modified)	Categories

Data from six studies (Popescu et al., 1991; Fossati et al., 2004; Martinez-Aran et al., 2004; Glahn et al., 2006; Taylor Tavares et al., 2007; Holmes et al., 2008) was included in the analysis of *working memory monitoring*. The pooled estimate of effect size was d= -0.682 (95%CI= -0.489 to -0.875; χ^2=48.0, p<0.0001). As the majority of data was from either the digit span (reverse) or SWM tests, these were further examined in separate analyses. For Reverse digit span, the pooled estimate of effect size was d= -0.932 (95%CI= -0.630 to -1.234; χ^2=36.5, p<0.0001) and for SWM was d= -0.410 (95%CI= -0.131 to -0.690; χ^2=8.3, *p*=0.004).

For the analysis of *Set shifting/ rule formation and reversal*, data from seven studies was included (Sweeney et al., 2000; Basso et al., 2002; Fossati et al., 2004; Martinez-Aran et al., 2004; Rubinsztein et al., 2006; Taylor Tavares et al., 2007; Roiser et al., 2009), with 2 studies providing information from two tests (Sweeney et al., 2000; Martinez-Aran et al., 2004). The pooled estimate of effect size was d= -0.416 (95%CI= -0.245 to -0.587; χ^2=22.8, p<0.0001). A number of these studies used the WCST and IDED tasks and therefore the pooled effect in these measures was examined separately. The pooled estimate of the effect size was d= -0.300 (95%CI = -0.093 to -0.506; χ^2= 8.1, p<0.005).

For the analysis of *Planning, reasoning and strategy*, data from six studies was included (Sweeney et al., 2000; Dixon et al., 2004; Rubinsztein et al., 2006; Taylor Tavares et al., 2007; Holmes et al., 2008; Roiser et al., 2009), with one study providing data from 2 tests (Sweeney et al., 2000) and one providing two patient samples against the same control group (Holmes et al., 2008). The pooled estimate of the effect size was d= -0.359 (95%CI= -0.184 to -0.533; χ^2=16.2, p<0.0001). In this analysis, most contrasts came from the SWM strategy score and therefore this analysis was repeated including only this data, producing a pooled estimated effect size of d= -0.335 (95%CI= -0.128 to -0.542; χ^2=10.0, *p*=0.0015).

For the analysis of *Inhibition*, data from five studies was included (Deptula et al., 1991; Popescu et al., 1991; Dixon et al., 2004; Martinez-Aran et al., 2004; Kerr et al., 2005), with two providing data from two tests each (Dixon et al., 2004; Martinez-Aran et al., 2004). The pooled estimate of effect size d= -1.074 (95%CI =-0.844 to -1.305; χ^2=83.6, p<0.0001). As the majority of measures within this domain are from either the Stroop colour-word test or the corrected interference effect, the analysis was repeated including only these measures. The pooled estimate of effect size was d= -1.044 (95%CI= -0.769 to -1.319; χ^2=55.4, p<0.0001).

For the analysis of *Fluency*, data from four studies was included (Neu et al., 2001; Basso et al., 2002; Dixon et al., 2004; Martinez-Aran et al., 2004), with two providing data from two tests each (Dixon et al., 2004; Martinez-Aran et al., 2004). The pooled estimate of effect size for these measures was d= -0.843 (95%CI= -0.616 to -1.069; χ^2= 53.3, p<0.0001). The fluency tests used can be separated into those assessing phonological fluency (Basso et al., 2002; Dixon et al., 2004; Martinez-Aran et al., 2004) and those assessing semantic fluency (Neu et al., 2001; Dixon et al., 2004; Martinez-Aran et al., 2004). When analysed separately, the pooled estimate of effect for category fluency was d= -0.803 (95%CI= -0.490 to -1.116; χ^2=25.3, p<0.0001) and for phonological fluency was d= -0.887 (95%CI= -0.559 to -1.215; χ^2=28.1, p<0.0001).

1.3.4 Immediate (short-term) memory impairment

In one of the more homogeneous neuropsychological domains, immediate (or short-term) memory has been assessed using the digit span and CANTAB spatial span (SSP) tasks. These tasks involve the temporary maintenance of verbal/phonological and spatial information respectively and accord to current models of working memory e.g. (Baddeley & Hitch, 1974).

Table 1-4. Summary of studies examining immediate memory tasks in bipolar depression

Paper	Study		BD			Con			TEST	Outcome measure
		Cohen's d	n	mean	s.d.	n	mean	s.d.		
25	Martinez-Aran, A., et al. 2004	-0.736*	30	5.3	1.5	30	6.3	1.2	Digit span	Forwards
11	Glahn, D. C., et al 2006	-0.090*	11	7.52	1.3	32	8.50	1.9	Digit span	Forwards
		-0.652*	15	8.10	1.8	32	8.50	1.9	Digit span	Forwards
10	Fossati, P., et al. 2004	-0.909*	17	5.8	1.1	49	6.8	1.1	Digit span	Forwards
30	Roiser, J. P., et al. 2009	-0.387*	49	6.1	1.7	55	6.7	1.4	SSP (CANTAB)	Forwards
36	Sweeney, J. A., et al. 2000	-0.694*	21	5.67	1.43	51	6.58	1.3	SSP (CANTAB)	Forwards
37	Taylor Tavares, J. V., et al 2007	-0.387*	17	6.19	1.36	25	6.75	1.5	SSP (CANTAB)	Forwards

Three studies included measures of forwards digit span (Fossati et al., 2004; Martinez-Aran et al., 2004; Glahn et al., 2006) and spatial span (Sweeney et al., 2000; Taylor Tavares et al., 2007; Roiser et al., 2009). The study by Glahn and colleagues examined patients with and without a history of psychosis and therefore there are 4 outcome measures for digit span, although the same control group is used in both comparisons.

For digit span (forwards) or *immediate verbal memory* the pooled estimate of the effect size is d= -0.622 (95%CI= -0.326 to - 0.917; χ^2 =17.0, p<0.0001).

For the spatial span task, or *immediate spatial memory,* the pooled estimate of the effect size is d= -0.470 (95%CI= -0.192 to - 0.748; χ^2 =17.0, *p*=0.0009).

1.3.5 Verbal memory impairments

The majority of tasks within this domain broadly follow the same format, with subjects being required to retain and recall or recognise verbal information either immediately or after a delay. Variations in procedures come in the number of items to be remembered, mode/frequency of presentation, and potential outcome measure. The Rey-Auditory Verbal Learning Test (Rey-AVLT) and California Verbal Learning Test (CVLT) have been used most frequently and require the participant to recall 15 or 16 words read by the experimenter over multiple presentations. After a delay period there are also measures of delayed recall and recognition (with the CVLT, some of these are cued or free). The remaining tasks differ slightly from these.

The *selective reminding task* also tests episodic memory; here participants search an A4 format card containing four verbal items, pointing to and naming each item (e.g. dentist) when its category cue (e.g. profession) is given verbally. After the search, cued recall is immediately tested and if one or more names is not recalled, the card is presented again for naming followed by cued recall, until all names are retrieved. This continues until all 16 items are identified and retrieved. The study phase is followed by three tests trials of free recall preceded by 20 s of interference by counting backward. Participants recall as many items as

possible and then are cued (using the category) for recall of the remainder. Following the last free and cued recall trials, a yes–no recognition memory for the items is administered.

Deptula's verbal memory test assesses recall and recognition of 20 words (all four-legged animals) which are read to the participant and repeated back to ensure attention. Recall is assessed over four trials of 150 seconds duration – only those items not recalled are subsequently read to the participant. On the second and fourth trials, forced choice recognition is also assessed.

Table 1-5. Summary of studies examining verbal memory tasks in bipolar depression

Paper	Study		BD			Con			TEST	Outcome measure
		Cohen's d	n	mean	s.d.	n	mean	s.d.		
1	Basso, M. R., et al. 2002	-2.072	25	83.44	6.11	34	92.25	2.01	CVLT	% savings
		-3.871	25	-0.7	0.32	34	0.39	0.25	CVLT	acquisition slope
		-1.176	25	-1.96	1.87	34	-0.29	0.97	CVLT	delayed recall (cued)
25	Martinez-Aran, A., et al. 2004	-1.335	30	9.5	3.1	30	13.2	2.4	CVLT	delayed recall (cued)
1	Basso, M. R., et al. 2002	-1.115*	25	-2.09	1.93	34	-0.41	1.1	CVLT	delayed recall (free)
25	Martinez-Aran, A., et al. 2004	-1.311*	30	8.6	3.1	30	12.6	3	CVLT	delayed recall (free)
1	Basso, M. R., et al. 2002	-0.908	25	-0.7	1.01	34	-0.06	0.34	CVLT	discrimination index
		-0.992	25	-0.75	1.42	34	0.35	0.81	CVLT	recall consistency
		-1.097*	25	-1.78	1.56	34	-0.29	1.19	CVLT	recognition hits
25	Martinez-Aran, A., et al. 2004	-1.031*	30	13.1	2.1	30	14.9	1.3	CVLT	recognition hits
1	Basso, M. R., et al. 2002	-0.687	25	-0.43	1.23	34	0.32	0.98	CVLT	semantic cluster ratio
		0.177	25	0.08	0.95	34	-0.08	0.87	CVLT	serial cluster ratio
		-0.862	25	-1.39	1.77	34	-0.21	0.98	CVLT	short recall (cued)
25	Martinez-Aran, A., et al. 2004	-1.085	30	9.9	3	30	12.8	2.3	CVLT	short recall (cued)
1	Basso, M. R., et al. 2002	-1.004	25	-1.87	1.89	34	-0.38	1.1	CVLT	short recall (free)
25	Martinez-Aran, A., et al. 2004	-1.138	30	7.9	3.3	30	11.6	3.2	CVLT	short recall (free)
1	Basso, M. R., et al. 2002	-1.240*	25	29.7	15.56	34	47	12.65	CVLT	total recall
25	Martinez-Aran, A., et al. 2004	-1.116*	30	43.4	10.1	30	54.4	9.6	CVLT	total recall
1	Basso, M. R., et al. 2002	-0.473*	25	-1.17	0.98	34	-0.62	1.28	CVLT	trial 1
		-1.449	25	-2.26	1.83	34	-0.21	1.01	CVLT	trial 5

Paper	Study		BD			Con			TEST	Outcome measure
		Cohen's d	n	mean	s.d.	n	mean	s.d.		
38	Wolfe, J., et al. 1987	-2.255	12	-	-	20	-	-	Rey-AVLT	Correct (A1 to 5)
27	Neu, P., et al. 2001	-0.656*	24	6.25	1.67	62	7.27	1.51	Rey-AVLT	Correct (A1)
		-0.500	24	8.58	2.48	62	9.62	1.91	Rey-AVLT	Correct (A2)
38	Wolfe, J., et al. 1987	-1.633	12	-	-	20	-	-	Rey-AVLT	Correct (A5 minus A1)
27	Neu, P., et al. 2001	-0.563*	24	6	2.84	62	7.38	2.29	Rey-AVLT	Correct (delayed)
38	Wolfe, J., et al. 1987	-2.082	12	-	-	20	-	-	Rey-AVLT	Recognition (% correct)
10	Fossati, P., et al. 2004	-0.805*	18	8.5	2.91	88	10.3	2.08	Selective reminding	Free recall 1
		-0.500	18	10.9	2.15	88	12	2.21	Selective reminding	Free recall 2
		-0.551	18	12.1	2.3	88	13.2	1.93	Selective reminding	Free recall 3
		-0.727*	18	31.6	6.4	88	35.6	5.31	Selective reminding	Sum of Free recall
		0.000	18	15.2	0.89	88	15.2	1.17	Selective reminding	Total recall 1
		0.000	18	15.6	0.77	88	15.6	0.78	Selective reminding	Total recall 2
		-0.285	18	15.8	0.51	88	15.9	0.31	Selective reminding	Total recall 3
		-0.684	18	42.2	18.58	80	54.9	18.58	Selective reminding	Index CFR%
7	Deptula, D., et al. 1991	-0.463*	8	54.7	7.5	14	60	13.1	Verbal recall	Recall
		-0.192	8	2.8	3.5	14	2.2	2.9	Verbal recall	Intrusions (%)
		-0.494*	8	2.73	0.99	14	3.1	0.58	Verbal recognition	Recognition
		-0.871	8	13.1	11.6	14	5.9	5.7	Verbal recognition	False positives (%)
25	Martinez-Aran, A., et al. 2004	-1.210*	30	49	8.9	30	59.3	8.1	WMS-R logical memory	Immediate recall
		-1.482*	30	44.6	7.9	30	55.8	7.2	WMS-R logical memory	Delayed recall

Finally, the last test employed is the Logical Memory subtest of the Wechsler Memory Scale, which assesses immediate and delayed recall of a word passage.

A number of effect sizes can be extracted from the available data. A *total immediate free-recall* measure is reported in 5 studies (Wolfe et al., 1987; Deptula et al., 1991; Basso et al., 2002; Fossati et al., 2004; Martinez-Aran et al., 2004), although there is insufficient data to include the Wolfe data in the pooled analysis. The pooled estimate of this effect size is d= -0.995 (95%CI = -0.736 to -1.254; χ^2=56.8, p<0.0001). It was noted that all tasks with the exception of logical memory involve the acquisition of individual words rather than a structure paragraph therefore the analysis was repeated with this measure removed and the effect size was only marginally different (d= -0.938, 95%CI = -0.645 to -1.232; χ^2=39.3, p<0.0001).

Some studies also included a measure of *initial immediate free-recall* i.e. recall performance after the first, single presentation of the to-be-remembered list (Neu et al., 2001; Basso et al., 2002; Fossati et al., 2004). The pooled estimate of effect size for this measure was d= -0.644 (95%CI= -0.352 to -0.937; χ^2=18.6, p<0.0001).

Several studies also included measures of longer-term verbal memory through the inclusion of *delayed recall* (Neu et al., 2001; Basso et al., 2002; Martinez-Aran et al., 2004)and/or *delayed recognition* (Deptula et al., 1991; Basso et al., 2002; Martinez-Aran et al., 2004). The pooled estimate of effect size for delayed recall was d= -1.064 (95%CI= -0.796 to -1.332; χ^2=60.6, p<0.0001) and for recognition d= -0.957 (95%CI= -0.604 to -1.310; χ^2=28.3, p<0.0001).

1.3.6 Visuo-spatial memory impairments

The majority of tasks included in this domain are from the CANTAB test battery. These can be broadly divided into those requiring the immediate or delayed recognition of visual patterns or spatial locations. Most follow a forced-choice recognition format. Within the visual domain, a number of studies have used the available variants of the simultaneous and delayed match-to-sample tests. The *Match to Sample - simultaneous* (MTS) version requires subjects to select the correct pattern from one, two, four, or eight peripheral patterns that matches the pattern shown simultaneously at the centre of the screen. The *Simultaneous/Delayed Match to Sample* (S/DMTS) test requires participants to select from a choice of 4 stimuli the one that correctly matches a target stimulus. This is done either simultaneously, or after the target has disappeared from the screen for delays of 0 sec, 4 sec or 12 sec. Other CANTAB tests employed include *Paired Associative Learning* (PAL) in which subjects are sequentially shown between one and eight patterns in an equally-spaced 'circle' around the screen. After a brief delay, the individual patterns are presented in the centre of the screen and the participant must indicate the peripheral location where it was first presented. This can therefore be thought of as an object-location binding task. Also, *Pattern and Spatial Recognition* (SREC and PREC) tasks, where participants are presented with a series of patterns or spatial locations and must indicate the correct item or position from a choice of two, after a brief delay.

The remaining tasks used include, the visual memory and reproduction sub-tests from the Wechsler Memory Scale, a non-verbal recall and recognition task (Deptula et al., 1991) and a novel Spatial Delayed Response Task (SDRT) in which participants view a target array of 1, 3, or 5 yellow circles, positioned pseudo-randomly around a central fixation, and after a fixed delay, are required to indicate whether a single green circle is in the same position as one of the target circles.

Table 1-6. Summary of studies examining visuo-spatial memory tasks in bipolar depression

Paper	Study		BD			Con			TEST	Outcome measure
		Cohen's d	n	mean	s.d.	n	mean	s.d.		
36	Sweeney, J. A., et al. 2000	-0.599*	21	95	5.5	51	97.2	2.6	MTS	Percent correct
		0.406*	21	1400	803	51	1656	546	MTS	Response latency (msec)
30	Roiser, J. P., et al. 2009	-0.093*	49	94	14.6	55	95.4	15.3	SMTS (CANTAB)	Percent correct
31	Rubinsztein, J. S., et al. 2006	0.000*	24	0.97	0.0	26	0.97	0.051	SMTS (CANTAB)	Proportion correct
36	Sweeney, J. A., et al. 2000	-0.690*	21	3340	1046	51	2834	561	SMTS (CANTAB)	Response latency (msec)
31	Rubinsztein, J. S., et al. 2006	-0.803*	24	3730	1151.3	26	2920	856.6	SMTS (CANTAB)	Response latency (msec)
30	Roiser, J. P., et al. 2009	-0.123	49	86.9	13.6	55	88.6	14.1	S/DMTS (CANTAB)	Delayed 0 sec % correct
		-0.266	49	85.4	17.4	55	89.4	12.6	S/DMTS (CANTAB)	4 sec % correct
		-0.306	49	77.3	21.9	55	82.9	14.3	S/DMTS (CANTAB)	12 sec % correct
37	Taylor Tavares, J. V., et al 2007	-0.311	17	0.849	0.12	25	0.885	0.11	S/DMTS (CANTAB)	Proportion correct
31	Rubinsztein, J. S., et al. 2006	-1.359	24	0.73	??	26	0.83	0.10	S/DMTS (CANTAB)	Proportion correct
30	Roiser, J. P., et al. 2009	-0.385	49	0.1	0.14	55	0.05	0.12	S/DMTS (CANTAB)	Prob error following error
31	Rubinsztein, J. S., et al. 2006	-0.583*	24	3952	1273.7	26	3324	856.6	S/DMTS (CANTAB)	Response latency (msec)
30	Roiser, J. P., et al. 2009	0.266*	49	3747	1233	55	4046	1020	S/DMTS (CANTAB)	Latency (msec)
37	Taylor Tavares, J. V., et al 2007	-0.202*	17	3869	1847	25	3601	815	S/DMTS (CANTAB)	Latency (msec)
36	Sweeney, J. A., et al. 2000	0.113*	21	3460	742	51	3549	808	S/DMTS (CANTAB)	Latency all delays (msec)
36	Sweeney, J. A., et al. 2000	-0.488*	21	22	40	51	10	14	PAL (CANTAB)	Total errors (adjusted)
27	Neu, P., et al. 2001	-0.171*	24	8.91	3.36	62	9.48	3.31	WMS	Correct score
25	Martinez-Aran, A., et al. 2004	-1.449	30	51.2	11.3	30	65.1	7.5	WMS-R (visual reproduction)	Delayed recall
		-1.156*	30	55.7	10.7	30	66.2	7.1	WMS-R (visual reproduction)	Immediate recall

Paper	Study		BD			Con			TEST	Outcome measure
		Cohen's d	n	mean	s.d.	n	mean	s.d.		
14	Holmes, M. K., et al. 2008	0.143*	32	1905	600	52	1998	682	PREC (CANTAB)	Latency
		-0.340*	33	2206	473	52	1998	682	PREC (CANTAB)	Latency
30	Roiser, J. P., et al. 2009	0.192*	49	1853	434	55	1957	622	PREC (CANTAB)	Latency
31	Rubinsztein, J. S., et al. 2006	-0.529*	24	2617	1361.9	26	2118.5	2.45	PREC (CANTAB)	Latency
14	Holmes, M. K., et al. 2008	-0.107*	32	91.5	8.4	52	92.3	6.9	PREC (CANTAB)	Percent correct
		-0.117*	33	91.4	8.8	52	92.3	6.9	PREC (CANTAB)	Percent correct
30	Roiser, J. P., et al. 2009	-0.208*	49	89.6	10.3	55	91.5	8	PREC (CANTAB)	Percent correct
36	Sweeney, J. A., et al. 2000	-0.232*	21	88.7	8.9	51	90.6	7.9	PREC (CANTAB)	Percent correct
31	Rubinsztein, J. S., et al. 2006	-0.006*	24	0.8	12.7	26	0.87	8.67	PREC (CANTAB)	Proportion correct
37	Taylor Tavares, J. V., et al 2007	-0.552*	17	0.869	0.1	25	0.924	0.1	PREC (CANTAB)	Proportion correct
7	Deptula, D., et al. 1991	-2.335	8	17.2	7.4	14	5.7	2.8	Non-verbal (recall)	Intrusions (%)
		-1.092	8	31.9	16	14	15.4	14.6	Non-verbal (recognition)	False positives (%)
		-2.860*	8	12	15	14	52	13	Non-verbal (recall)	Recall
		-1.145*	8	1.77	0.69	14	2.46	0.55	Non-verbal (recognition)	Recognition
11	Glahn, D. C., et al 2006	-0.033*	15	39.53	3.2	32	39.72	3.5	Spatial DRT	delayed response
		-0.856*	11	36.73	2.1	32	39.72	3.5	Spatial DRT	delayed response
30	Roiser, J. P., et al. 2009	0.249	49	1972	670	55	2144	710	SREC (CANTAB)	Latency
31	Rubinsztein, J. S., et al. 2006	-0.219	24	2615	926	26	2428	780.1	SREC (CANTAB)	Latency
30	Roiser, J. P., et al. 2009	0.016*	49	81.5	11.6	55	81.3	12.7	SREC (CANTAB)	Percent correct
36	Sweeney, J. A., et al. 2000	-0.395*	21	77.4	15.4	51	82.3	11	SREC (CANTAB)	Percent correct
31	Rubinsztein, J. S., et al. 2006	-0.007*	24	0.7	12	26	0.79	12.2	SREC (CANTAB)	Proportion correct
37	Taylor Tavares, J. V., et al 2007	-0.279*	17	0.806	0.12	25	0.846	0.155	SREC (CANTAB)	Proportion correct

Within the domain of visuo-spatial memory, the relative homogeneity of test selection allows a number of pooled effect sizes to be extracted from the available data. *Simultaneous matching-to-sample tests* (i.e. where the test stimuli and the target appear together and must be matched) provide both accuracy (Sweeney et al., 2000; Rubinsztein et al., 2006; Roiser et al., 2009) and latency (Sweeney et al., 2000; Rubinsztein et al., 2006) measures. The pooled estimate of effect size for accuracy was d= -0.207 (95%CI=0.063 to -0.477; χ^2=2.26, *p*=0.133) and for latency d= -0.321 (95%CI= -0.012 to -0.629; χ^2=4.16, *p*=0.041).

For the *delayed match-to-sample (DMTS)* test, although accuracy was reported in several tests there was no common outcome measure used in all. The pooled estimate of effect size for latency (Sweeney et al., 2000; Rubinsztein et al., 2006; Taylor Tavares et al., 2007; Roiser et al., 2009) was d= -0.007 (95%CI=0.241 to -0.255; χ^2=0.003, *p*=0.955).

Tests assessing memory for visual/pattern stimuli fitted into those that required participants to recall or reproduce this information from memory (Deptula et al., 1991; Neu et al., 2001; Martinez-Aran et al., 2004), and those that required forced-choice recognition of the correct item (Deptula et al., 1991; Sweeney et al., 2000; Rubinsztein et al., 2006; Taylor Tavares et al., 2007; Holmes et al., 2008; Roiser et al., 2009). The pooled estimate of effect size for *visual memory (immediate recall)* was d= -0.769 (95%CI= -0.427 to -1.111; χ^2=19.4, p<0.0001) and for *visual memory (recognition)* was d= -0.167 (95%CI=0.022 to -0.356; χ^2=3.0, *p*=0.083).

It is clear that the task used in one study (Deptula et al., 1991) produced a more pronounced effect that the other studies in this analysis which all used the CANTAB Pattern Recognition test (PRec), therefore the analysis was repeated using these studies and produced a pooled estimated effect size of d= -0.128 (95%CI=0.065 to -0.321; χ^2=1.69, *p*=0.193) for accuracy. The

pooled estimate of the latency values from this measure was d= -0.068 (95%CI=0.155 to -0.291; χ^2=0.358, *p*=0.550).

Within the domain of *spatial memory (recognition)*, all tasks included some measure of accuracy (Sweeney et al., 2000; Glahn et al., 2006; Rubinsztein et al., 2006; Taylor Tavares et al., 2007; Roiser et al., 2009). The pooled estimate of effect size was d= -0.222 (95%CI= -0.022 to -0.423; χ^2=4.71, *p*=0.030).

1.3.7 (Psycho)motor speed impairments

Broadly, the tasks employed within this domain can be categorized as those that test fine motor skills or dexterity and those that assess reaction time to respond to an event (sometimes with a decision making component).

The Trail Making Test (TMT) part A (Reiten, 1958) assesses simple motor dexterity, requiring participants to joint up the numbers 1 to 25 in order, as quickly as possible (Part B is used to assess set-shifting; executive function). The Finger Tapping Tests require participants to 'tap' as many times as possible within a given time period while in the Grooved Pegboard, participants must place as many small pegs in key-like slots as quickly as possible.

Within the reaction time tests, the Five Stage Reaction Time task (CANTAB) assesses psychomotor speed. Subjects are asked to hold down a press pad and release it and touch a yellow dot on the screen as soon as it appears. Simple reaction time requires participants to respond as quickly as possible to a target, while choice reaction times are similar but responses are conditional on the nature of the probe (i.e. response to red not blue lights) or a different key must be pressed in response to the different probes (Choice Reaction Time; double).

Other tests included, counting tasks and rate of articulation where participants are required to count as quickly as possible, or without pausing.

Table 1-7. Summary of studies examining (psycho)motor tasks in bipolar depression

Paper	Study		BD			Con			TEST	Outcome measure
		Cohen's d	n	mean	s.d.	n	mean	s.d.		
5	Burdick, K. E., et al. 2009	-1.160*	24	704.5	189.4	24	524.7	110.2	Choice RT	(log) ms
28	Popescu, C., et al. 1991	-1.464	24	3.2	2.5	30	0.57	0.94	Choice RT (double)	Errors
		-0.710	24	51.07	14.29	30	43.67	5.74	Choice RT (double)	RT (latency)
		-0.281	24	1.31	1.89	30	0.9	0.99	Choice RT (single)	Errors
		-0.686*	24	44.53	10.05	30	38.98	6.11	Choice RT (single)	RT (latency)
5	Burdick, K. E., et al. 2009	-0.200*	24	52.5	10.4	24	54.4	8.5	Finger tapping test	total
36	Sweeney, J. A., et al. 2000	0.247	21	570	124	51	602	132	Five Stage RT	Movement time (msec)
		-0.367	21	385	117	51	354	67	Five Stage RT	RT (msec)
1	Basso, M. R., et al. 2002	-1.401*	25	83.12	21.97	34	61.36	8.12	Grooved pegboard	Dominant hand
		-1.645	25	3.04	1.54	34	0.88	1.12	Grooved pegboard	Impaired scores
		-1.164	25	90.54	24.99	34	68.76	12.25	Grooved pegboard	Non-dominant hand
5	Burdick, K. E., et al. 2009	-0.729*	24	71.8	16.1	24	63	5.7	Grooved pegboard	time to complete
28	Popescu, C., et al. 1991	-0.414*	24	57.51	11.55	30	53.27	9.05	Motor RT	RT (latency)
5	Burdick, K. E., et al. 2009	-0.726*	24	525	144.2	24	431.4	111.5	Simple RT	(log) ms
28	Popescu, C., et al. 1991	-0.521*	24	26.88	8.57	30	23.76	2.44	Simple RT	RT (latency)
		-1.024	24	90.88	9.73	30	102.1	11.84	Tapping speed	LH; 15sec
		-0.951	24	174.92	18.77	30	193.5	20.12	Tapping speed	LH; 30sec
		-0.724	24	100.92	9.04	30	109.7	14.1	Tapping speed	RH; 15sec
		-0.782*	24	193.88	17.74	30	211.1	24.9	Tapping speed	RH; 30sec
1	Basso, M. R., et al. 2002	-1.057*	25	31.04	9.29	34	24.3	2.74	Trails	A
25	Martinez-Aran, A., et al. 2004	-1.062*	30	51.2	25.4	30	30.1	12	Trails	A
27	Neu, P., et al. 2001	-0.549*	24	43.03	17.71	62	34.08	15.75	Trails	A
28	Popescu, C., et al. 1991	-0.053	24	4.77	1.92	30	4.67	1.88	Counting task	1 to 10
		0.037	24	11.19	4.52	30	11.37	5.04	Counting task	1 to 20
8	Dixon, T., et al. 2004	-0.711	15	0.3	0.2	30	0.2	0.1	Rate of articulation	mean sec per word
34	Schneider, J. J., et al. 2008	-1.626	32	4.8	1.7	28	8.6	2.9	WAIS-III (executive)	coding
		-1.086	32	6.4	2.3	28	9	2.5	WAIS-III (executive)	symbol search

For the *fine motor and dexterity analysis*, data were pooled from 3 types of test (Trails A, tapping tests, and grooved pegboard) taken from five studies (Popescu et al., 1991; Neu et al., 2001; Basso et al., 2002; Martinez-Aran et al., 2004; Burdick et al., 2009). The pooled estimate of the effect size was d= -0.794 (95%CI= -0.588 to -1.001; χ^2=56.9, p<0.0001).

As there were a number of different studies that used the Trails A, the pooled effect size for this test was examined separately and resulted in an estimated effect size of d= -0.845 (95%CI= -0.545 to -1.145; χ^2=30.5, p<0.0001).

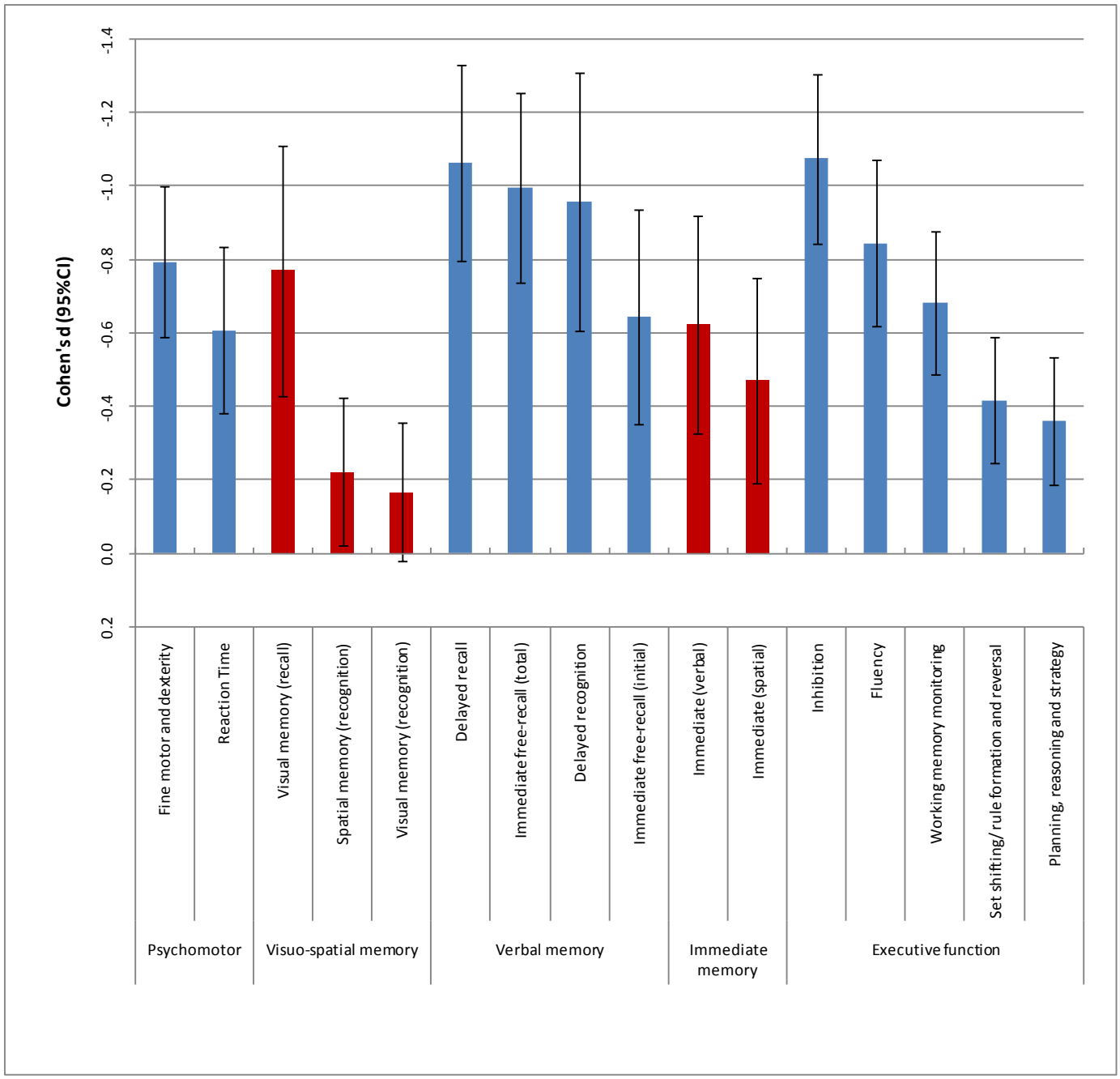
Of the *reaction time* tests, these are taken from three studies only (Popescu et al., 1991; Sweeney et al., 2000; Burdick et al., 2009) therefore multiple measures/tests are from the same samples. The pooled estimate of effect size is d= -0.607 (95%CI= -0.380 to -0.833; χ^2=27.6, p<0.0001).

1.3.8 Summary of effect size analysis

Examining the methodological issues first, as discussed in the above sections, there are several instances where multiple 'valid' patient samples come from the same study and are therefore compared against the same control group. This occurs, for example, when comparing groups with and without psychosis or on and off medication. Using the same controls may reduce variance within the domains covered by these studies. Also, the effect sizes generated are simple estimated effects and are not adjusted for demographic differences (through covariates) as they are in many analyses. The estimated effects are consequently susceptible to inflating the true effect size and should be viewed as a method of general comparison only. Finally, some domains or processes examined are covered by a more limited and homogeneous set of tests than others (e.g. immediate memory using the digit span or spatial span tests). Consequently there is likely to be less variance in these areas than in instances where a more broad definition was used to pool effects.

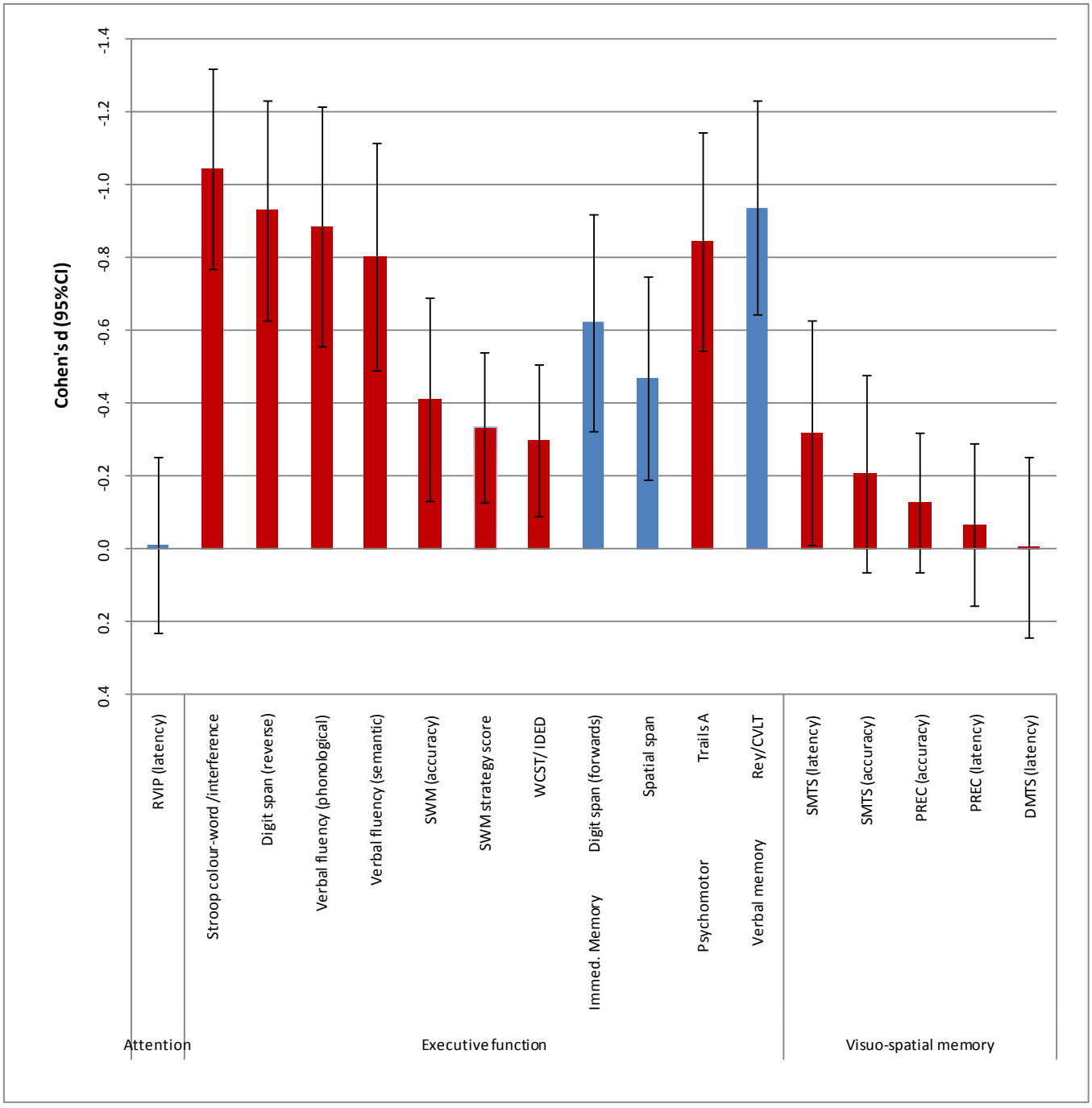
Overall, the analysis of pooled effect sizes yielded results largely consistent with the existing non-meta analytic reviews of the literature (Quraishi & Frangou, 2002; Malhi et al., 2004; Savitz et al., 2005) although some differences were noted. For example, it had been proposed that within the domain of verbal fluency, only semantic fluency was impaired in bipolar depression (Malhi et al., 2004) whereas the current analysis found similar effect sizes for both phonological and semantic fluency. The largest effect sizes were observed in the domains of verbal memory, as well as aspects of executive functioning (inhibition, fluency and working memory monitoring), psychomotor speed (fine motor and dexterity), and visuo-spatial memory (immediate visual recall). However, focusing on processes rather than domains it is clear that within memory functioning, measures that are assessed by recall rather than recognition and/or assess delayed rather than immediate recollection, yield the largest effects.

Figure 1-1. Summary of pooled effect sizes (with lower and upper 95%CI) for primary neuropsychological processes in bipolar depression



Different colouring of bars is for clarity, to denote separate neuropsychological domains (i.e. psychomotor, visuo-spatial etc.).

Figure 1-2. Summary of pooled effect sizes (with lower and upper 95%CI) for individual tests in bipolar depression



The size of effects are summarised in Figure 1-1, where within the visuo-spatial memory tests, visual recall produces a pooled effect size around d= -0.8, with visual and spatial recognition being around d= -0.2 (unfortunately no study included a measure of spatial memory recall; See Chapter 5 of the thesis for a test that includes this measure). Within the verbal memory, immediate and delayed recall as well as delayed recognition produce effect sizes around d= -1.0. It is worth noting that the executive domain includes a wider range of estimated effect sizes for different processes, likely reflecting the complexity and heterogeneity of measures included in this domain. Again, it should be highlighted that some of these processes are pooled from very similar or the same measures e.g. inhibition is purely from Stroop colour-word or interference scores, while others are drawn from a variety of measures.

Previous reviews of this area have produced conclusions attesting to the general, broad profile of memory impairment in bipolar depression: *"…deficits are likely the consequence of reduced cognitive effort or inefficient encoding and retrieval strategies, which result in poor free recall"* (Malhi et al., 2004). A similar conclusion can be drawn from the present review of the literature. This has relevance for the subsequent course of this thesis and leads directly to the selection tests in Chapter 2; the initial examination of neuropsychological functioning in bipolar depressed patients. It also leads to the need in Chapter 4 to assess the factor structure of neuropsychological functioning in a larger sample of participants, developing understanding of how these processes relate to one another. Prior to this, the next section examines the background literature on HPA axis dysfunction in mood disorder and its inclusion in this thesis as a modulator of neuropsychological functioning.

1.4 HPA axis dysregulation in mood disorder

In this section, a brief overview of the HPA axis is first presented followed by those studies which have examined dysregulation in mood disorders. Although pertinent literature of depressive disorders in general will be examined, there will be a focus on bipolar depression.

1.4.1 The HPA axis

The HPA axis is one of the major hormonal systems mediating physical and psychological stress responses. When activated, neurones in the paraventricular nucleus of the hypothalamus secrete corticotropin-releasing hormone which is transported via the hypothalamo-pituitary portal circulation to the anterior pituitary where adrenocorticotropic hormone (ACTH) is secreted through stimulation of pituitary corticotrophs. ACTH then stimulates the adrenal cortex to secrete glucocorticoids: corticosterone in rats and cortisol in humans (Feldman et al., 1995; Berne & Levy, 1998).

1.4.1.1 Cortisol and corticosteroid receptors

Under basal conditions, cortisol secretion exhibits a 24 hour circadian rhythm in which concentrations are highest at waking and slowly decline to a nocturnal trough (Weitzman et al., 1971). As with many hormones it is released in a pulsatile manner throughout this cycle (Young et al., 2004a). A great deal of individual variation exists in the secretion of both ACTH and cortisol, but spontaneously occurring cortisol peaks are preceded by increases in ACTH levels, although secretion of the two hormones are not quantitatively linked throughout the day (Follenius et al., 1987). Indeed, analysis of ultradian variations within healthy individuals has shown a predominant periodicity in the oscillations of both hormones of between 55 and 140 minutes for ACTH and 95 and 180 minutes for cortisol, indicating that, on occasion, a single cortisol peak may be initiated by two ACTH peaks (Follenius et al., 1987). Levels also

appear to exhibit seasonal variation with plasma cortisol being higher in winter, but overall cortisol production rate reduced (Walker et al., 1997; King et al., 2000; Hansen et al., 2001).

Cortisol is involved in the regulation of fat, protein and carbohydrate metabolism, electrolyte balance, body water distribution, blood pressure and immunosuppressant anti-inflammatory action (Berne & Levy, 1998). As discussed it is also a key regulator of the physiological stress response, through negative-feedback actions via corticosteroid receptors. Two distinct corticosteroid receptor subtypes have been identified; the mineralocorticoid receptor (MR; Type I) and the glucocorticoid receptor (GR; Type II). Both receptor types have been implicated in mediating glucocorticoid feedback (Reul & de Kloet, 1985), however there are several differences in the distribution, occupancy and binding properties of the two receptors that affects their role physiologically. The MR is highly expressed in the limbic system whereas the GR is ubiquitous, being present in both subcortical and cortical structures, with a preferential distribution in the prefrontal cortex (Patel et al., 2000).

Glucocorticoids bind to the MR with around a 6- to 10- fold greater affinity than to GR (de Kloet et al., 1999). Consequently, at basal levels near complete occupation of MRs occurs. GRs are minimally occupied at this point and only during times of high cortisol secretion, such as the circadian peak or during stress, do MRs become saturated and GR occupancy increases (Reul & de Kloet, 1985; de Kloet & Reul, 1987). A growing body of evidence indicates that alterations in HPA axis function may be a core feature of mood disorders and may exert significant causal and exacerbating effects on symptoms and neuropsychological functioning (for a review see Anacker et al.; Gallagher et al., 2009).

1.4.1.2 DHEA

Recently there has been increased interest in the role of other adrenal steroids such as dehydroepiandrosterone (DHEA) which, in its sulfated form (DHEAS) is the most abundant adrenal steroid in humans (Morfin, 2002). DHEA is a naturally occurring excitatory neuroactive steroid (or neurosteroid: a term first proposed by Baulieu and colleagues in 1981 (Baulieu, 1981) that applies to the steroids, the accumulation of which occurs in the nervous system independently, at least in part, of supply by the steroidogenic endocrine glands and which can be synthesized de novo in the nervous system (Baulieu & Robel, 1998)). The neurosteroidogenesis in the brain is independent of the peripheral production; brain DHEAS was not influenced by adrenal stimulation or inhibition with adrenocorticotropic hormone (ACTH) or dexamethasone, respectively, and increased 2 days after the stressful event of adrenalectomy and orchiectomy (Corpechot et al., 1981). DHEA is a substrate for androstenedione and testosterone synthesis and may have a role as an adrenal androgen (Gurnell & Chatterjee, 2001). DHEA serves as a precursor of androstenedione, testosterone, as well as of approximately 50% of androgens in adult men, 75% of active estrogens in premenopausal women, and 100% of active estrogens after menopause (Regelson & Kalimi, 1994).

As with cortisol, DHEA levels have been shown to exhibit seasonal variation (Garde et al., 2000) although other studies have not found such changes and results seem far less consistent (Bjornerem et al., 2006; Brambilla et al., 2007). The diurnal rhythm of DHEA also appears to be less pronounced than that of cortisol (Hucklebridge et al., 2005). Neurosteroids display multiple effects on the central nervous system (CNS) and may act as potential signalling molecules for neocortical organization during neuronal development (Baulieu & Robel, 1996; Mao & Barger, 1998). In particular, neurosteroids can interact with various neurotransmitter systems to promote neuronal remodelling; they regulate growth of neurons, enhance

myelinisation and synaptogenesis in the CNS, affect synaptic functioning, and show neuroprotective properties (Friess et al., 2000; Wen et al., 2001; Johnson et al., 2002). Furthermore, these neurosteroids have been found in the mammalian brain at considerably higher concentrations than typically detected in serum or plasma (Corpechot et al., 1981; Baulieu, 1997). There is evidence that neurosteroids may be involved in the vulnerability to developing neuropsychiatric disorders such as dementia, mood disorders, substance abuse and others (for reviews see; (Epperson et al., 1999; Sundstrom Poromaa et al., 2003; Eser et al., 2006; Girdler & Klatzkin, 2007; Ritsner et al., 2008).

The precise mechanism of action of DHEA in the brain is less well known although it has been shown to have actions on membrane-bound receptors and is a gamma-aminobutyric acid type A ($GABA_A$) receptor antagonist (Hansen et al., 1999) as well as a sigma-1 receptor agonist (Maurice et al., 1999; Maurice et al., 2006). Recently it has been confirmed that neuroactive steroids (pregnenolone, DHEA, DHEAS, allopregnanolone) are present in human post-mortem brain tissue at physiologically relevant concentrations in the nanomolar range and that levels of pregnenolone and DHEA in posterior cingulate and parietal cortex are higher in subjects with schizophrenia and bipolar disorder compared to control subjects (Marx et al., 2006). However, in addition to these neurosteroid properties, it is the putative role of DHEA(S) as a functional antagonist of the actions of cortisol which have generated most interest in the study of patients with psychiatric illness.

1.4.1.3 Methods of assessment

A number of methods are available for the assessment of basal steroid levels in humans. For example, for small, highly lipid-soluble molecules (such as cortisol) the unbound hormone can pass easily through the membranes of nucleated cells permitting ‘free’ steroid levels to appear

in bodily fluids (Kirschbaum & Hellhammer, 2000). Levels can reliably be measured in urine, plasma and saliva with each having potential strengths and weaknesses (Levine et al., 2007).

1.4.1.3.1 Urine

Urinary cortisol excretion results from glomerular filtration and is a useful index of integrated 24-hour plasma free cortisol (Levine et al., 2007), but steroid output over any fixed period of time can be reliably assessed (Callies et al., 2000). Similarly reliable measurements of DHEA as well as many other steroid metabolites can be achieved (Poor et al., 2004).

1.4.1.3.2 Saliva

Saliva sampling has certain advantages over plasma sampling, especially in patients whose HPA axes may be sensitive to stressful interventions such as venepuncture (Kirschbaum & Hellhammer, 1989, 1994, 2000; Lac, 2001). Due to the relatively small samples required to obtain steroid measurements, sampling can also be performed relatively frequently if necessary and can allow the circadian profile to be determined (Lac, 2001). The analysis of the area-under-the-curve provides an estimate of the overall hormonal secretion over 24-hour and – although not as precise – is more convenient than 24-hour urine collection.

Importantly, several studies have examined the relationship between steroid levels in saliva compared to those in plasma. Cortisol and DHEA levels measured in saliva closely agree with free levels in the blood, due to the fact that cellular access and entry to the oral cavity are by a method of passive diffusion and therefore independent of saliva flow-rate and transport mechanisms (Vining et al., 1983; Granger et al., 1999; Kirschbaum & Hellhammer, 2000).

However, it should be noted that the method of collection of saliva can have effects on the accuracy of this relationship (Granger et al., 1999; Shirtcliff et al., 2001; Gallagher et al., 2006).

1.4.1.3.3 Plasma

A relationship between central and peripheral steroid levels has been established (Carroll et al., 1976b). However, there are factors that can result in variability in this relationship, for example degree of blood brain permeability, activity of the multidrug-resistance gene type-1 P-glycoprotein, and steroid metabolising enzymes within the brain, such as 11β-hydroxysteroid dehydrogenase type 1 (Pearson et al., 2010). Guazzo et al (1996) measured plasma and CSF levels of cortisol and DHEA(S) in a group of 62 subjects aged 3 to 85 years. Significant correlations in steroid-free subjects were observed between blood and CSF levels for DHEA ($r = 0.65$) and DHEAS ($r = 0.88$) but not for cortisol ($r = 0.26$). However, in the case of cortisol, there appeared to be some evidence of two distinct populations diverging at blood concentrations of 300 to 400 nmol, with a strong relationship evident in one of these. Also, a strong relationship between CSF and blood levels emerged in the in a sub-group of participants on exogenous steroid administration (Guazzo et al., 1996).

Many studies have examined aspects of HPA axis dysfunction using the variety of methodologies described above. These are briefly discussed in the following section before reviewing the evidence for the utility of examining the ratio of adrenal steroid secretion.

1.4.2 Peripheral and basal abnormalities in mood disorders

The first systematic studies of the abnormalities in steroid hormone secretion in psychiatric illnesses were carried out by Board and colleagues over half a century ago (Board et al., 1956; Board et al., 1957). These initial findings were subsequently replicated by other groups and extended to show that levels reduced as patients recovered (Gibbons & McHugh, 1962;

Gibbons, 1964, 1966). Many studies have replicated these findings using a variety of methodologies and collection methods.

1.4.2.1 Urinary levels

In major depression, elevated urinary cortisol levels have been reported in many studies (e.g. Carroll et al., 1976a; Kathol et al., 1989; Maes et al., 1998; Scott & Dinan, 1998) and may persist in recovery in some patients (Kathol, 1985). Although this pattern may not be evident in some sub-groups of patients and may even reverse with age/ illness chronicity (Oldehinkel et al., 2001). Urinary DHEA levels have similarly been found to be elevated (Tollefson et al., 1990). The psychotic sub-type of unipolar and bipolar disorders also appears to be associated with higher urinary cortisol levels (Wedekind et al., 2007).

More recently, comprehensive analysis of multiple urinary steroid metabolites in medication-free patients with recurrent unipolar major depression revealed sex differences in some metabolites (Poor et al., 2004). In male patients (compared to male controls) levels of DHEA, as well as tetrahydrocorticosterone (THB), allo-THB, beta-cortolone (beta-CL) were found to be significantly decreased. However, in female patients, DHEA levels did not significantly differ from their respective control group, although cortisol and allo-THB levels were significantly elevated, and etiocholanolone and beta-CL levels were significantly decreased (Poor et al., 2004). Relationships between the ratio of cortisol and DHEA and their metabolites have also been examined in MDD in relation to symptom severity, with 11-beta-hydroxysteroid dehydrogenase (HSD) being correlated with severity in women, and 17-beta-HSD being positively correlated with severity in women but negatively correlated in men (Raven & Taylor, 1998).

Few studies have directly compared groups of patients with different diagnoses, although of those that have it has been found that 24-hour urinary cortisol levels were higher in affective disorders compared with schizophrenia (Diebold et al., 1981; Yehuda et al., 1993).

1.4.2.2 Saliva

As discussed earlier, because of the ease with which samples can be collected, many studies have examined steroid levels in saliva (those assessing both cortisol and DHEA or the ratio in the same samples are presented in more detail subsequently in this chapter).

Recently there has been interest in the measurement of cortisol levels in saliva for the first hour after waking when cortisol levels are known to sharply rise. Several studies have demonstrated that clear abnormalities can be observed in patients with mood disorders. Unmedicated depressed patients have been found to secrete up to 25% more cortisol in the first hour after waking than control subjects (Bhagwagar et al., 2005). This increased cortisol awakening response (CAR) has been found to persist in remitted depressed patients (Bhagwagar et al., 2003). Similarly, increased CAR has been observed in clinically well patients with bipolar disorder, with normal DST responses (Deshauer et al., 2003) and recently in young high-risk subjects who had never personally suffered from depression but who had a biological parent with a history of major depression (Mannie et al., 2007). However, it should be noted that the interpretation of results of the CAR can be complex as some authors have highlighted that a blunted CAR can also be assumed to be abnormal (Aas et al., 2010).

1.4.2.3 Plasma and CSF

Plasma sampling has often been adopted to take point-estimates of adrenal steroid secretion although as discussed, the pulsatile nature of release limits interpretation of findings. Of greater interest are those studies that have sampled at multiple time-points throughout the day to accurately profile the pattern of steroid secretion. One method proposed by Halbreich – sampling cortisol levels in plasma every half-hour (from 1pm to 4pm) correlates well with 24-hr cortisol levels and provides a reliable estimate of hypercortisolaemia (Halbreich et al., 1982).

Other studies have used even more extended sampling periods. Wong and colleagues (Wong et al., 2000) performed a comprehensive assessment of plasma and CSF steroids every 30 minutes over 30 hours in medication-free melancholic MDD patients. ACTH levels were not significantly different from healthy controls. However, cortisol levels were significantly elevated as was the cortisol/ACTH ratio suggesting a relatively greater plasma cortisol response to a given simultaneous level of plasma ACTH (Wong et al., 2000). CSF cortisol levels have been shown to be elevated in both unipolar and bipolar disorder, with even greater levels evident in patients with psychotic features (Carroll et al., 1976b).

Those studies assessing DHEA(S) levels in affective disorders present a somewhat mixed picture (for an overview see (van Broekhoven & Verkes, 2003)), although many have looked at the effect on depressive symptoms rather than a clear diagnosis of mood disorder. Other differences likely arise due to methodological factors or through assessment of steroids in isolation rather than considering the relationship with other adrenal steroids (see below).

1.4.2.4 The Cortisol/DHEA ratio

It has been suggested that the assessment of cortisol or DHEA(S) alone may not be as informative as calculating the ratio of the two steroids – the cortisol to DHEA or DHEAS/cortisol molar ratio (Hechter et al., 1997). The notion is that DHEA(S) may maintain cortisol homeostasis by acting as a cortisol antagonist, particularly during periods of prolonged glucocorticoid hyperactivity. Several lines of evidence have shown that a variety of stressors result in a shift in the balance of cortisol and DHEA(S), in that there is an increase in cortisol synthesis and a decrease in androgen synthesis. In critical illness it has been demonstrated that not only do plasma levels of cortisol increase and DHEA decrease, but sensitivity of both to ACTH-stimulation is also correspondingly altered (Parker et al., 1985). Similarly, during acute psychological stress, stimulation of adrenal steroid release is accompanied by a shift towards DHEA release (Oberbeck et al., 1998). This has also led to the recognition of the potent antiglucocorticoid properties of DHEA(S) (Kalimi et al., 1994) (and its active metabolites, see (Muller et al., 2006)). In animals it has been demonstrated that DHEA protects hippocampal neurons against neurotoxin-induced cell death, possibly by decreasing nuclear GR levels (Cardounel et al., 1999). DHEA(S) has also been shown to inhibit glucocorticoid-induced enzyme activity (Browne et al., 1992). In healthy humans, acute administration of DHEA has been shown to rapidly reduce circulating cortisol levels (Wolf et al., 1997) while reduction in 24 hour levels have been demonstrated with longer treatment trials in healthy older subjects (Kroboth et al., 2003).

Since DHEA levels appear to have regulatory effects on glucocorticoid action in the brain, it has been argued that the ratio of cortisol to DHEA most accurately reflects the degree of 'functional' hypercortisolaemia (Goodyer et al., 1998; Wolkowitz et al., 2001; Gallagher & Young, 2002). Together, these studies highlight the importance of considering the somewhat

symbiotic relationship between cortisol and DHEA(S) and suggests that examination of each in isolation may fail to be as informative as assessment of the ratio of the two.

An extensive series of longitudinal studies examining risk factors for the development of mood disorders in adolescents by Goodyer and colleagues (Goodyer et al., 1996; Herbert et al., 1996; Goodyer et al., 1998; Goodyer et al., 2000a, 2000b; Goodyer et al., 2001a; Goodyer et al., 2001b; Goodyer et al., 2003) showed that the secretion of adrenal steroids is altered and of predictive utility. In saliva samples collected over 48 hours it was found that elevated evening cortisol and lower morning DHEA secretion were significantly, and independently, associated with major depression (Goodyer et al., 1996). Different patterns of adrenal steroid secretion were associated with co-morbidity (Herbert et al., 1996).

Young and colleagues (Young et al., 2002) assessed salivary cortisol to DHEA molar ratios over 2 consecutive days (at 8 am and 8 pm) in 44 medication-free major depressed patients compared to their matched controls. All patients were drug-free for at least 6 weeks although most were entirely medication-naïve (n = 26/44) and of the 18 who had previously received psychotropic medication, the time drug-free ranged from 6 to 336 weeks (median = 48 weeks). Depressive symptom scores in the patient group ranged from 15 to 30 (mean = 21) on the Hamilton Depression Rating Scale. Thirty patients (68%) were experiencing their first episode of depression. Although cortisol levels were elevated and DHEA levels decreased in the patient group, neither difference reached statistical significance however the molar cortisol/DHEA ratio was significantly elevated. It should be noted that saliva was collected using a salivette device which, as discussed previously, studies have shown can affect the accuracy of DHEA measurement (Granger et al., 1999; Shirtcliff et al., 2001; Gallagher et al., 2006).

In a comparison of depressed and remitted patients with major depressive disorder (the majority of whom were taking antidepressant medication) and matched controls, Michael and colleagues reported that salivary cortisol to DHEA ratios were significantly elevated, both at 8 am and 8 pm, compared to remitted patients and healthy controls who did not significantly differ. Furthermore, in a post hoc analysis, taking the 85th percentile morning (8 am) cortisol/DHEA ratio of the control group as a cut-off, 82.5% of the depressed group had cortisol/DHEA ratios that were equal to or greater than this value, while this occurred in only 15% of healthy controls (Michael et al., 2000).

Using an intensive sampling methodology, Heuser and colleagues collected blood samples every 30 minutes over 24 hours in 26 depressed patients and 33 controls for assessment of cortisol and DHEA levels. Mean cortisol and DHEA levels, and minimum DHEA level was found to be elevated over the 24 hour period compared to controls (Heuser et al., 1998). An elevation in both cortisol and DHEA was also observed in a smaller group of female depressed patients (Weber et al., 2000). However the cortisol to DHEA molar ratio was not calculated in either study. Interestingly, it was noted that these finding differed from a smaller earlier study which sampled blood at single time-points where cortisol levels were significantly elevated while DHEA did not differ (Osran et al., 1993). Here the cortisol to DHEA ratio was also calculated and was found to be elevated in the morning (8 am) samples but not at 4 pm (Osran et al., 1993). Elevated cortisol levels and cortisol to DHEA ratios have also been found in un-medicated female MDD patients with co-morbid borderline personality disorder (Kahl et al., 2006) and in elderly depressed patients (Ferrari et al., 2004) although ageing itself is noted to significantly reduce DHEAS secretion (Ferrari et al., 2001a; Ferrari et al., 2001b).

More discrepant results have been found in studies adopting single plasma-sampling methodology. In medication-free subjects (>4 weeks), Scott and colleagues (Scott et al., 1999)

found evidence of increased ratios of both cortisol to DHEA, and cortisol to DHEAS. When assessed individually, neither cortisol nor DHEA levels differed significantly from controls although DHEAS levels were lower in the patients (Scott et al., 1999). However, lower levels of both DHEA and cortisol with no difference in DHEAS have also been reported (Jozuka et al., 2003) although here the ratio was not calculated.

Other studies measuring the sulfated form have found elevated salivary DHEAS levels, even in the absence of abnormal cortisol levels in medicated patients with MDD. Although the sample size was somewhat modest, discriminant analysis indicated that 77% of subjects could be correctly classified by evening DHEAS levels (Assies et al., 2004). In a preliminary study, Takebayashi and colleagues found DHEAS and cortisol levels to be significantly elevated compared to controls in plasma samples taken at baseline in an outpatient sample (aged <45 years). Following treatment, DHEAS had significantly decreased. There were no differences in the DHEAS to cortisol ratio of patients and controls at any point (Takebayashi et al., 1998). In one study of older depressed patients (>60 years) compared with matched controls, no differences in DHEA(S) to cortisol ratios were reported (Fabian et al., 2001).

Very little work has been carried out on assessing cortisol to DHEA ratios in patients with bipolar disorder. Using a repeated plasma sampling protocol, hypercortisolaemia has been observed in bipolar patients (with depressive symptoms) compared with controls (chapter 2), without alteration in DHEA levels or cortisol to DHEA molar ratio (Gallagher et al., 2007)

1.4.3 Activating/integrated tests

The most sensitive tests of HPA axis function, however, are 'activating' tests whereby neuroendocrine responses are measured following pharmacological challenge. These are preferred not only because of their increased sensitivity, but because they elucidate

functional changes in the HPA axis at the receptor level (Watson et al., 2006a). The GR agonist dexamethasone has been used widely to examine HPA axis negative feedback integrity (Rush et al., 1996). An abnormal (non-suppressed) cortisol response to dexamethasone administration has been described in patients with mood disorder (Rush et al., 1996) and may be more pronounced in those with psychotic features (Duval et al., 2000). The combined dexamethasone/corticotropin releasing hormone (dex/CRH) test (Heuser et al., 1994) is also abnormal during relapse (Heuser et al., 1994; Modell et al., 1997) and persists in recovery, particularly in bipolar disorder (Rybakowski & Twardowska, 1999; Watson et al., 2004; Watson et al., 2007). Furthermore, corticosteroid receptor abnormalities have been observed in post-mortem studies which show evidence of regionally-specific changes in MR and GR mRNA expression in post-mortem brain tissue samples from patients with mood disorders (Knable et al., 2001; Webster et al., 2002; Lopez et al., 2003).

It has been suggested that raised cortisol is a marker of prognosis and that HPA axis dysfunction and persistent hypercortisolaemia are likely to identify those patients who are either not improving or are likely to be vulnerable to relapse. In a study of depressed patients treated with the SSRI fluoxetine (Young et al., 2004b), non-responders showed abnormal HPA axis reactivity, whilst responders did not differ from healthy controls. A meta-analysis of the dexamethasone suppression test (DST) as a predictor of treatment outcome concluded that although DST status at baseline was not predictive of response to antidepressant treatment, persistent non-suppression DST after treatment was associated with high risk of early relapse and poor outcome after discharge from hospital (Ribeiro et al., 1993). Several studies have examined neuroendocrine responses to the dex/CRH test in depression and have found relationships with relapse or treatment response (Ising et al., 2007). For example, it has been shown that in clinically remitted major depression, post-treatment responses to the dex/CRH were significantly higher among patients who relapsed (Appelhof et al., 2006; Aubry et al.,

2007). It has been argued that early improvement, early treatment response and beneficial treatment outcome are associated with a lower HPA axis activity (assessed using the dex/CRH) and that, in the longer-term, HPA axis dysregulation increases in parallel with the number of previous episodes (Hatzinger et al., 2002).

1.5 Is there any evidence of a link between HPA axis dysregulation and neuropsychological impairment?

In the following section evidence for the direct relationship between cortisol and cognition is examined. First the work from animal studies is discussed before focussing on the extension of this to healthy human subjects and clinical conditions, particularly mood disorders. It should be noted that this will broadly examine the topic as more detailed reviews are available elsewhere, such as the acute effects of glucocorticoids on cognition (Lupien & McEwen, 1997; Het et al., 2005), the effects of stress on cognitive function (Sauro et al., 2003), and the modulatory effects of emotional content on memory (Roozendaal et al., 2008).

1.5.1 Animal work

Much of the work from animal models brings together two lines of study – the known effects of corticosteroids on hippocampal function and the role of the hippocampal formation in learning and memory. Lupien and McEwen comprehensively reviewed the literature on the acute effects of corticosteroids on memory in animals and humans (Lupien & McEwen, 1997). The overall pattern of results highlight that there are clear dose-related effects that elicit either facilitation in memory, as doses rise from sub-optimal to optimal levels, or impairment at higher doses and therefore show an inverted "U"-shape relationship. The role of the

specific receptor sub-types in different stages of memory formation can be separated from the effects of glucocorticoids have on general arousal. The model proposes that the dose-response relationship emerges because of differential activation of MR or GR, especially in the hippocampus, with MR being involved in the processes of sensory integration and GR with acquisition and consolidation of the memory trace (Lupien & McEwen, 1997). It is also noteworthy that some of these effects may occur via interaction with neurotransmitter/ hormone complexes, such as sex hormones (Symonds et al., 2004; Kuhlmann & Wolf, 2005) and noradrenergic (Quirarte et al., 1997; Roozendaal et al., 2006) or serotonergic (McAllister-Williams et al., 1998; Porter et al., 2002; Pariante et al., 2004a; Porter et al., 2004) systems. For example, chronic elevation of glucocorticoid levels, by corticosterone administration or stress, causes functional desensitization of the 5-HT_{1A} autoreceptor (Lanfumey et al., 1999; Fairchild et al., 2003). Functional 5-HT_{1A} autoreceptor desensitization also occurs when corticosterone rhythm is flattened at a level around the mid-diurnal level (Gartside et al., 2003; Leitch et al., 2003). Recently it has been noted that endocannabinoids in the amygdala enhance memory consolidation and that cannabinoid-receptor activity within this brain region may be required for enabling glucocorticoid effects on such memory processes (Campolongo et al., 2009).

1.5.2 Healthy human work

In humans it has been suggested that a distinction be made between the effect of corticosteroids on general arousal or attention and their effect on specific memory processes, paralleling that described in the animal literature (Lupien & McEwen, 1997). Several studies have demonstrated verbal declarative memory deficits following administration of hydrocortisone (Wolkowitz et al., 1990; Newcomer et al., 1999). Those that have examined different stages of information processing have argued that this is specifically an effect on memory retrieval (de Quervain et al., 2000; de Quervain et al., 2003). This appears discrepant

from the findings in rodents where the effects are on acquisition and consolidation. However de Quervain and colleagues suggest that it is the delay interval (24 hours) that is crucial to this difference. Although memory is impaired in humans by pre-learning administration of cortisol (e.g. Newcomer et al., 1999) glucocorticoid levels may remain elevated at the time of testing in these experiments. Thus it is possible that such results actually reflect impaired memory retrieval rather than altered memory acquisition or consolidation. A recent review has also highlighted the effect of diurnal changes in cortisol levels on the memory effects of glucocorticoids (Het et al., 2005). It is worth noting that all of these studies have used declarative verbal recall to assess performance, although it has been proposed that verbal working memory may be more sensitive than declarative memory to the acute effects of glucocorticoids (Lupien et al., 1999). There is also evidence that after sub-chronic doses of hydrocortisone, spatial working memory is also impaired (Young et al., 1999). This introduces interesting opportunities to parallel the work in animals which has a focus on spatial memory and the hippocampus. Including emotional content into the word lists can further affect the pattern of impairment and facilitation following cortisol elevation (for more detailed primary data and reviews see for example Wolf et al., 2004; Smeets et al., 2008; Wolf, 2008).

A meta-analysis has suggested a more complex picture of the effects of stress on memory (i.e. when memory is assessed following acute laboratory stress or long term exposure to rising basal levels of glucocorticoids) compared with that found after pharmacological manipulation (Sauro et al., 2003). This highlights the difficulty of generalising results across differing methods of HPA axis manipulation, and an even greater degree of complexity when attempting to apply these models of glucocorticoid-cognition interactions to clinical conditions in which cortisol levels/receptor dysfunction may be present over long periods of time. Broadly, clinical studies can be separated into those that have looked for direct

associations between cortisol levels and cognitive performance and those that have used markers of the general magnitude of HPA disturbance and neuropsychological impairment.

1.5.3 Clinical conditions

That hypersecretion of cortisol may be causative in the development of depression is suggested by findings in patients with Cushing's disease/syndrome (CD/CS). Typically, individuals have plasma cortisol levels that are three-fold those of healthy subjects, and within this group the prevalence of depression is higher than in the normal population (Cohen, 1980; Kelly et al., 1983). Furthermore, depressive symptoms resolve on treatment of the primary endocrine disorder (Cohen, 1980; Kelly et al., 1983). Importantly, there is also clear impairment in neuropsychological functioning in these individuals. A number of studies have now demonstrated impairments in learning and memory, delayed recall, and visual-spatial ability (Whelan et al., 1980; Mauri et al., 1993; Forget et al., 2000; Starkman et al., 2001).

Particular focus has also been placed on the effects of hypercortisolaemia in CS/D on hippocampal structures, where volume reductions have been noted in a significant proportion of patients (Starkman et al., 1992). These reductions are noted to be strongly correlated with impairment on tasks of verbal learning and recall – tasks known to be sensitive to hippocampal complex/temporal lobe dysfunction – and also with the degree of hypercortisolaemia (Starkman et al., 1992). Multiple mechanisms by which glucocorticoids induce these morphological changes in the brain have been posited, including decreased glucose utilization, increased actions of excitatory amino acids, inhibition of long-term potentiation and decreased neurotrophic factors, and decreased neurogenesis (Patil et al., 2007). It should be noted that in CS/D, the temporal relationship between treatment and recovery of neuropsychological functioning is not always coincident (Forget et al., 2002).

Zobel and colleagues found that antidepressant treatment-associated changes in the cortisol response to the dex/CRH test in patients with major depression were correlated with improvements in working memory but not with improvements in episodic memory, sustained attention or global severity of symptoms (Zobel et al., 2004). This is consistent with the results of studies in healthy subjects discussed above (Lupien et al., 1999). However this finding has not yet been replicated. In contrast, Reppermund and colleagues assessed neuropsychological performance and administered the dex/CRH test to a group of 75 depressed inpatients of which 51 (68%) were in remission at the point of discharge. Despite a significant reduction of depressive symptoms between admission and discharge, high rates of neuropsychological impairment were still observed. Selective attention did improve in remitted and non-remitted patients, while speed of information processing improved only in those who had remitted. The cortisol response to the dex/CRH test decreased significantly only in remitted patients, but this was not correlated with neuropsychological performance. In non-remitted patients, severity of depression was significantly correlated with information processing while improvement in short-term memory was negatively associated with the cortisol response at discharge. Thus, it appears that HPA axis dysregulation and symptom severity have differential effects on verbal short term memory and speed of information processing (Reppermund et al., 2007).

The potential importance of the association between the consequences of HPA axis dysregulation and neuropsychological performance was indirectly illustrated recently in a study by Gorwood and colleagues exploring the hypothesis of the 'toxic' effects of depression on the hippocampus. Using verbal declarative memory (the delayed paragraph recall index from the Wechsler Memory Scale-Revised) as a surrogate marker of hippocampal function, 8,229 patients were assessed twice over an average 42 day period and Structural Equation Modelling used to assess the clinical and demographic factors predicting performance. At

presentation, current illness severity was an important determinant of performance, while previous depressive history (the number and length of past episodes) was not. At the follow-up after significant clinical response, the intensity of previous depressive history was more significant than current symptoms. Crucially, an inverse relationship was found between performance and recurrence whereby each additional episode (up to 4 episodes) impaired verbal declarative memory performance by 2% to 3% (Gorwood et al., 2008). The direct relationship of this finding to measures of HPA axis dysregulation and hypercortisolaemia requires further study.

Higher levels of morning salivary cortisol have been shown to be associated with post-encoding memory retrieval and storage deficits and executive dysfunction in major depression, in the absence of any relationship with symptoms (Egeland et al., 2005). This is consistent with the findings of Reppermund and colleagues described above. In psychotic major depression, relationships between elevated mean cortisol levels (1800 to 0100 hours) and poorer verbal memory and psychomotor speed performance have been found, while the cortisol slope over this period significantly correlated with both verbal memory and working memory (Gomez et al., 2006). In contrast, one recent study in first-episode psychosis, a blunted CAR response[3] was found to be was associated with a more severe deficit in verbal memory and processing speed (Aas et al., 2010). Inverse relationships between peripheral cortisol levels and general intellectual functioning, but not verbal declarative memory have also been reported in major depression (van Londen et al., 1998).

[3] It is noted by the authors that this blunted response is assumed to be abnormal.

1.6 General summary and implications for the current thesis

The preceding chapter outlines the literature surrounding the main research themes for the present thesis and the subsequent three empirical chapters. In this chapter a general introduction to the illness was presented (1.2) followed by a review of neuropsychological impairments in bipolar depression (1.3) and one of the potential neurobiological underpinnings of this phenomenon, that is, HPA axis dysfunction (1.4,1.5). As has been discussed, of the time that patients are symptomatic, the majority of this time is in depression and therefore this is the focus of the current research. When depressed, individuals present with a broad range of neuropsychological impairments, with the largest effect sizes seen in aspects of executive functioning, verbal and visual free-recall and delayed recall/ recognition, and psychomotor fine-motor control. A significant proportion of individuals also exhibit HPA axis dysfunction and consequent hypercortisolaemia, which animal and human studies have implicated in the modulation of neuropsychological functions. Exploring these features of the illness is not only important from a treatment perspective, but in bipolar disorder itself we may have a clinical condition in which we have the opportunity to examine the link between corticosteroids and memory.

Therefore, in Chapter 2, the first empirical study is presented which assesses neuropsychological functioning and hypercortisolaemia in individuals with bipolar disorder with depressive symptoms. In Chapter 3, the results of a study of the use of an antiglucocorticoid adjunctive treatment in the same participants are presented. The aim of this latter study is to examine the specific neuropsychological changes that occur as a result of modulating the HPA axis, specifically through GR antagonism and the potential amelioration of hypercortisolaemia.

Chapter II

Neuropsychological functioning and the HPA axis in bipolar depression

2. Neuropsychological functioning and the HPA axis in bipolar depression

2.1 Introduction to Chapter 2

As discussed previously (General introduction, section 1.3.2) there are relatively few studies which have specifically examined neuropsychological function in bipolar depression. The aim of this first empirical chapter is to present the results of the initial assessment of neuropsychological functioning in 20 patients with bipolar depression (compared to performance in 20 healthy controls) who subsequently entered the study of adjunctive antiglucocorticoid treatment.

The selection of neuropsychological tests was informed by the results of the review presented in Chapter 1.3.2 with the objective of broadly covering the domains described. Specifically, the intention was to include measures of executive function and attention, verbal and visuo-spatial memory and psychomotor speed. Many of the individual tests selected were characterised by large effect sizes in the pooled analyses and therefore increase statistical power in the present comparison. However, it is important to note that at this stage, the aim was to profile these functions broadly and not just to select those tests with the largest effects. To afford some protection against Type-I statistical error, a multivariate approach will be taken with the analysis and individual tests examined following significant global effects (see Stevens, 2002).

The neuroendocrine assessment was based on the method proposed by Halbreich of sampling cortisol levels in plasma, every half-hour, from 1pm to 4pm. This sampling period correlates well with 24-hr levels and provides a reliable estimate of hypercortisolaemia (Halbreich et al.,

1982). In addition, in a novel extension of this procedure, DHEA levels will also be measured in these samples to permit the estimation of cortisol-DHEA ratios (see General Introduction; 1.4.1.3).

2.2 Subjects and methods

2.2.1 Ethical approval

The study was approved by the Newcastle and North Tyneside LREC. After a full description of the study and all questions relating to the study had been answered, all subjects gave their written informed consent prior to inclusion. Subjects were free to leave the study at any point.

2.2.2 Subjects

2.2.2.1 Recruitment

Patients were recruited by opportunity sampling through referring clinicians who first assessed their suitability for the study, based on their current clinical state. Most patients were referred via a tertiary mood disorders service in the North East of England. Expenses incurred for travel and subsistence were reimbursed.

2.2.2.2 Matching of patient group to control reference sample

A normative sample of healthy control data, matching the demographic profile of the patient sample, was collated from a large database on file. To avoid selection bias, the SAS 'Match' algorithm was used (Kosanke & Bergstralh, 1995). The macro is used to match one or more controls (from a total of M) for each of N cases. The control selected for a particular case(i) is the control(j) closest to the case in terms of Dij, where Dij is the weighted sum of the absolute

differences between the case and control matching factors. A maximum accepted difference (DMAXK) is used to define the largest possible absolute differences compatible with a valid match. Cases are not matched to a control if any of the individual matching factor differences are >DMAXK. The algorithm also offers a 'greedy' or 'optimal' matching method – in the case of the former, once a match between case and control is made it is never broken; in the case of the latter, the 'PROC NETFLOW' command is used to find the set of matches that minimizes the sum of Dij over all possible sets of matches. The optimal method was selected. The matching variable used were age, sex and NART estimated IQ (given the weighting 2:2:1 in the algorithm). The tolerance of the match was ±5 yrs in age, exact match in sex, and ±16 NART points. Although this range of NART scores is higher than would be expected for a 'match', as the program is producing case-control matches, the matching of the overall group is very similar to the mean and variance of the patient sample (see Appendix 9.2).

2.2.2.3 Screening Assessment

Bipolar patients underwent both physical examination and psychiatric assessment prior to inclusion.

The physical examination was carried out to exclude significant medical illness, and past history of illness that may affect neuroendocrine or neurocognitive functioning. This included head-injury and neurological disorders. Further to this, some additional requirements pertaining to the antiglucocorticoid administration are described in chapter 3.

The psychiatric assessment was performed as a structured clinical interview by experienced psychiatrists. Bipolar patients were required to fulfil DSM-IV SCID (First et al., 1995) criteria for a major depressive episode (MDE) and all general inclusion/ exclusion criteria set out in Table

2-1**.** Control subjects were required to have no current or past history of psychiatric illness, and no first degree relative with current or past history. Additional inclusion/ exclusion criteria are also set out in Table 2-1**.**

During the screening interview, several rating scales were utilised to assess severity or subtype of depression in patients and to screen for depressive symptoms in controls. Descriptions of those used are presented below in Section 2.2.3.

2.2.2.4 Inclusion / Exclusion criteria

General inclusion/ exclusion criteria for the patient and control group are presented:

Table 2-1. Inclusion/ Exclusion criteria for patients and control group

BD patients	Control group
• Age 18 – 65 years	• Aged 18 – 65 years
• Fulfilling DSM-IV (SCID) criteria for bipolar disorder; current episode depressed (or with depressive symptoms)	• No history of depression or psychiatric illness
	• No first-degree relatives with a history of depression or psychiatric illness
	• Beck < 8
• No current alcohol dependence or abuse	• Alcohol intake $\leq$ 28 units per week (female $\leq$ 21 units)
• Physically healthy	• Physically healthy
• No ECT within the last 6 months	

2.2.3 Rating scales

BD subjects were assessed using all rating scales described. Control subjects completed the Beck Depression Inventory only.

2.2.3.1 Hamilton Depression Rating Scale (HAM-D)

The HAM-D is a frequently used depression rating scale in experimental and clinical research. One limitation however, is that the scale was devised for use only on patients already diagnosed as suffering from affective disorder of the depressive type and as such, it is not suitable for administration to control subjects.

The scale consists of 17 items (HAM-D_{17}), some defined in terms of a series of categories of increasing intensity and others by a number of equal-valued terms . Eight items are scored from 0 to 4 (0 = absent, 1 = mild, 2-3 = moderate, 4 = severe) and 9 scored from 0 to 2 (0 = absent, 1 = slight or doubtful, 2 = clearly present). A 21-item version is available (HAM-D_{21}) which includes 4 additional variables: diurnal variation, derealisation, paranoid symptoms and obsessional symptoms (see Appendix 9.3). Both versions were used in this study. The maximum score from the 17-item scale is 50, and from the 21-item the maximum is 62 (Hamilton, 1960).

2.2.3.2 Montgomery-Åsberg Depression Rating Scale (MADRS)

The MADRS was designed as a rating scale for depression which was especially sensitive to change, and therefore is of particular use in clinical trials or in smaller cohorts. Sensitivity and accuracy of change estimates were the major criteria for the inclusion of items. The scale is useful as it can be used with any time interval between ratings (Montgomery & Asberg, 1979).

The scale consists of ten items: apparent sadness, reported sadness, inner tension, reduced sleep, reduced appetite, concentration difficulties, lassitude, inability to feel, pessimistic thoughts, suicidal thoughts. These are rated by the clinician on a six point scale, with descriptors provided for the defined scale steps 0, 2, 4 and 6. The maximum score for the scale is 60 (see Appendix 9.4).

2.2.3.3 Beck Depression Inventory (BDI)

This self-report scale provides an assessment of the intensity of depressive symptoms. The scale is suitable for use in both subjects with a confirmed diagnosis of depression and those without.

The BDI consists of 21 categories, each one describing a specific behavioural manifestation of depression (see Appendix 9.5). For each category there is a graded series of 4 statements and each statement is assigned a score from 0 to 3. A description at the top of the BDI instructs the subject to read each statement in a category and to select the one which best represents how they have felt over the past week (Beck et al., 1961). The maximum score on the scale is 63.

2.2.3.4 Mania/hypomania

The use of the SCID in the screening assessment meant that only patients meeting criteria for bipolar depression were included in the study. However, to assess low level manic symptoms, the Young scale was also administered (Young et al., 1978).

2.2.4 Neuropsychological Tests

2.2.4.1 Design of the neurocognitive test battery

The neurocognitive test battery was designed to test several broad cognitive domains, across several modalities (visuo-spatial and verbal/auditory). Full descriptions of each test are provided in the following section.

Prior to administration of the main test battery, the NART (Nelson, 1982) was administered to estimate pre-morbid verbal IQ.

Table 2-2. Neurocognitive test battery

TEST	DOMAIN
1. NART	General screening (estimated verbal IQ)
2. Rey-AVLT	Learning and memory (verbal)
3. Vigil CPT	Sustained attention / Executive
4. Verbal fluency	Executive
5. Digit Span	Immediate memory / Executive
6. Rey-AVLT (Long-term)	Learning and memory (verbal)
7. SWM (CANTAB)	Executive
8. DSST	Psychomotor
9. PRec (CANTAB)	Learning and memory (visuo-spatial)
10. SRec (CANTAB)	Learning and memory (visuo-spatial)
11. SSp (CANTAB)	Immediate memory
12. Stroop	Executive

Computerised tests are shaded. Numbers represent order of administration.

2.2.4.2 Procedure

For the neurocognitive assessment, all subjects were tested at 1300h, with the testing taking approximately 75 minutes to complete. The neurocognitive test battery was designed to assess a broad range of cognitive domains and included pen-and-paper and computerised tests. Pen-and-paper tasks were administered according to standardised instructions (Lezak, 1995). Computerised tests were administered according to the CANTAB manual protocols, on a PC which was fitted with a colour touch-screen monitor fixed in a standardised position. Detailed descriptions of each test and details of their administration are given below.

2.2.4.3 General Screening Tests

2.2.4.3.1 National Adult Reading Test (NART) (Nelson, 1982)

The NART was originally designed as a method of estimating pre-morbid verbal IQ in subjects with dementing conditions (Lezak, 1995). By using phonetically-irregular words, it is not possible for subjects to 'sound out' a word that is not already known to them.

The version used in the present study consists of 50 such words (see Table 2-3) which the subject is required to pronounce correctly. The number of errors can then be converted to an estimate of pre-morbid verbal IQ. To ensure consistent administration of the test, a pronunciation guide is used by the experimenter in which the words are written using the International Phonetic Alphabet.

Table 2-3. Words from the National Adult Reading Test (NART)

Chord	Courteous	Hiatus	Facade	Gauche
Ache	Rarefy	Subtle	Zealot	Topiary
Depot	Equivocal	Procreate	Drachm	Leviathan
Aisle	Naïve	Gist	Aeon	Beatify
Bouquet	Catacomb	Gouge	Placebo	Prelate
Psalm	Gaoled	Superfluous	Abstemious	Sidereal
Capon	Thyme	Simile	Détente	Demesne
Deny	Heir	Banal	Idyll	Syncope
Nausea	Radix	Quadruped	Puerperal	Labile
Debt	Assignate	Cellist	Aver	Campanile

2.2.4.4 Immediate (short-term) memory

These tests are included separately as in theoretical terms they fit closely with current models of working memory architecture and are a good measure of phonological loop and visuo-spatial sketchpad capacity (Baddeley & Hitch, 1974).

2.2.4.4.1 Digit span (forwards)

It this test, the participant is read a series of numbers and they are asked to repeat them in the same order as given. The length of the number sequence increases until 2 incorrect responses at any given level. The maximum span attainable is 9.

2.2.4.4.2 Spatial span

This is a computerised test from the CANTAB battery and is analogous to the Corsi blocks test. It is the spatial equivalent of the digit span test (above). Participants view a fixed array of boxes on the screen and these change colour in a random order. Participants are then asked

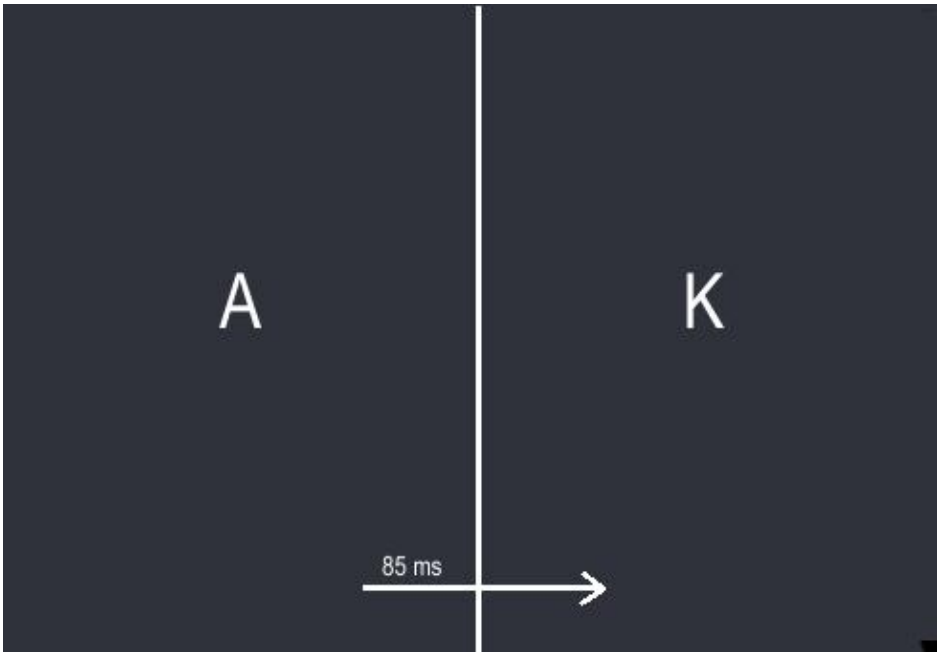
to tap out the same sequence on the computer touch-screen. The test terminates with three incorrect responses at any given level. The maximum attainable span is 9.

2.2.4.5 Attention and Executive Function

2.2.4.5.1 Vigil Continuous Performance Task

This is a continuous performance test that measures the ability to sustain attention over a period of time (Cegalis & Bowlin, 1991). It is a computerised test lasting 8 minutes in which single letters are flashed on the screen for 85 milliseconds (ms), with a gap of 915ms from one letter to the next, during which time the screen is blank. The subject is required to respond to the letter sequence of an 'A' followed by a 'K', and not to respond to any other stimuli. During the 8 minutes the subject is shown 480 trial stimuli (letters), among which there are 25 target stimuli ('A-K' sequences) in each 2 minute quarter of the test (100 in total). *Response latency* and errors of *omission* (where subjects fail to respond to an A-K sequence) and *commission* (where subjects respond to a stimulus other than the A-K sequence) are recorded.

Figure 2-1. Vigil Continuous Performance Test

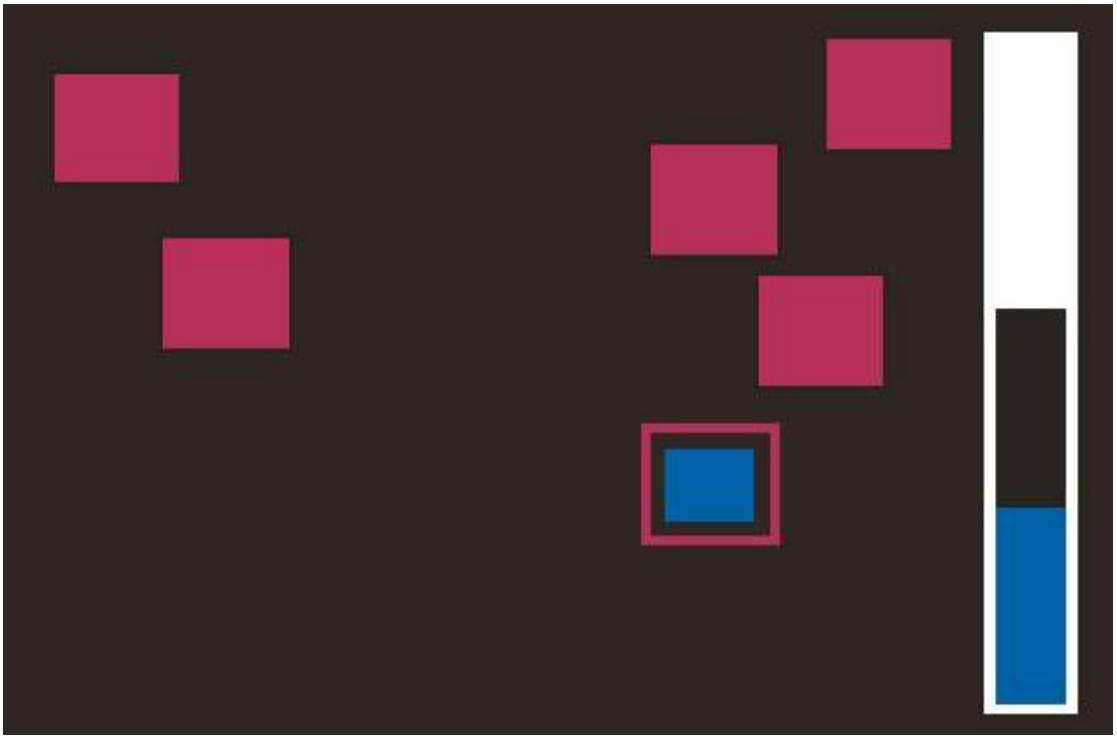


2.2.4.5.2 Spatial Working Memory (SWM)

This is a self-ordered search task which places demands on spatial working memory and executive function. Subjects must search through an increasing number of "boxes" (two, three, four, six, or eight) for a hidden token. Once a token is found, it will not appear in the same box again. Subjects must continue the search without returning to a box which has already contained a token. Accuracy is measured as the number of *between search errors* (the number of times boxes which have already contained tokens on previous trials are searched) and *within search errors* (the number of times boxes which have already been examined on the current trial are searched).

A *strategy score* is also computed, based on the use of a systematic search pattern on the 6 and 8 box problems. Higher scores indicate less use of the strategy.

Figure 2-2. CANTAB Spatial Working Memory test



2.2.4.5.3 Controlled Oral Word Association Test (Benton's FAS)

This is a test of verbal fluency which is sensitive to frontal lobe damage (Lezak, 1995). There are 3 trials, each lasting 60 seconds, in which subjects are required to list as many words as possible, beginning with the given letters – 'F', 'A' and 'S', excluding proper nouns, numbers or repetitions of the same word with a different suffix. Accuracy is measured as the overall number of legitimate correct words.

2.2.4.5.4 Stroop

There are many variants of the original Stroop test (for a review see MacLeod & MacDonald, 2000). The version used here has 2 parts to it: in the first, participants must read aloud the printed names of colours as quickly as possible. Time to complete is recorded. In the second part, participants are required to inhibit this response and instead state the colour of the ink that each word is incongruously printed in. Number correctly completed in 120 sec is recorded (Trenerry et al., 1989).

2.2.4.5.5 Digit span (reverse)

This is the second part of the span task described above (section 2.2.4.4.1). Administration is identical with the exception that participants must repeat back the sequence of digits in reverse order, thus placing greater demands on the online maintenance of working memory.

2.2.4.6 Psychomotor Performance

2.2.4.6.1 Digit Symbol Substitution Task (DSST)

This is a test of psychomotor speed and selective sustained attention from the Wechsler Adult Intelligence Scale-Revised (WAIS-R; (Wechsler, 1981)). The test consists of four rows of 25 blank squares, each with a random number between 1 and 9 associated with it. At the top of the page is a printed key in which each different number has been associated with an geometric symbol. Subjects are allowed to complete 7 samples which are not timed, and are then instructed to work as 'quickly and as accurately as possible' to work through each row/ square in order and draw the appropriate matched symbol in each of the squares given. The test is terminated after 90 seconds and the number correct in this time is recorded.

2.2.4.7 Learning and memory (verbal)

2.2.4.7.1 Rey – Auditory Verbal Learning Test (Rey – AVLT)

This test measures immediate memory, provides a learning curve, elicits retroactive and proactive interference tendencies and tendencies to confusion or confabulation on memory tasks and measures both short-term and longer-term retention following interpolated activity (Rey, 1964; Lezak, 1995).

It begins with a test of immediate word recall: for trial I, the examiner reads a list of 15 words (List A) at the rate of one per second. The subject is instructed to repeat back as many of the words as possible, in any order. This is then repeated a total of five times, with recall of the list recorded after each one (Trial I to V).

Immediately after trial V, the experimenter then reads out a different list of words (List B), instructing the subject to again repeat back as many words as possible. Following the B-list trial, the examiner asks the patient to recall as many words from the A-list as possible (trial VI) without further presentation of that list.

After a 30 minute delay, recall of List A is again tested (trial VII). A recognition trial is normally given whenever a subject's delayed recall is less than 13 words, however all subjects in the present study completed this trial. In testing recognition, the administrator asks the patient to identify as many words as possible from List A when shown a list of 50 words containing all the items from both the A and B lists as well as words that are semantically associated or phonemically similar to words on lists A or B.

The number of words correctly recalled or recognised are recorded. As performance on the final 2 recall trials of List A depends upon how well the words were initially learned, these scores are calculated as a percentage of the maximum score from the first 5 recalls. Interference indices can also be derived: *Proactive inhibition*, where previously learned material interferes with the acquisition of new material, can be calculated by subtracting the first recall of List A from recall of List B and *Retroactive inhibition*, where material learned after the to-be-remembered list interferes with subsequent recall of that list, can be calculated by subtracting the fifth recall of List A from the sixth.

2.2.4.8 Learning and Memory (visuo-spatial)

All tests of visuo-spatial learning and memory were from the CAmbridge Neuropsychological Test Automated Battery (CANTAB).

2.2.4.8.1 Pattern Recognition (PRec)

This is a test of visual recognition memory. A total of 24 visual patterns are presented in the centre of the screen (in 2 sets of 12) for three seconds each. These patterns are designed so that they cannot easily be give verbal labels. In the recognition phase, after a delay of five seconds, pairs of patterns are presented. Subjects are required to choose between a pattern they have already seen and a novel pattern. The test patterns are presented in the reverse order to the original order of presentation. Immediate auditory and visual feedback (a green tick or a red cross) is provided for accuracy of response. Accuracy measured as the total percentage correct. Speed of response is also recorded, measured as the latency between the pair of patterns appearing on the screen and the subject's response.

2.2.4.8.2 Spatial Recognition (SRec)

This is a test of spatial recognition memory. The subject is presented with a white square that moves in sequence to five different locations on the screen. Each remains on the screen for a total of 3 seconds. In the recognition phase, after a 5 second delay, the subject sees a series of five pairs of squares, one of which is in a location *not* seen in the presentation phase. The subject must touch the box which *is* in exactly the same location as one of those in the initial presentation sequence. As with the pattern recognition test, locations are tested in the reverse of the presentation order. A total of four trials are completed. Total percentage correct and response latency are recorded.

Figure 2.1: CANTAB Pattern Recognition test

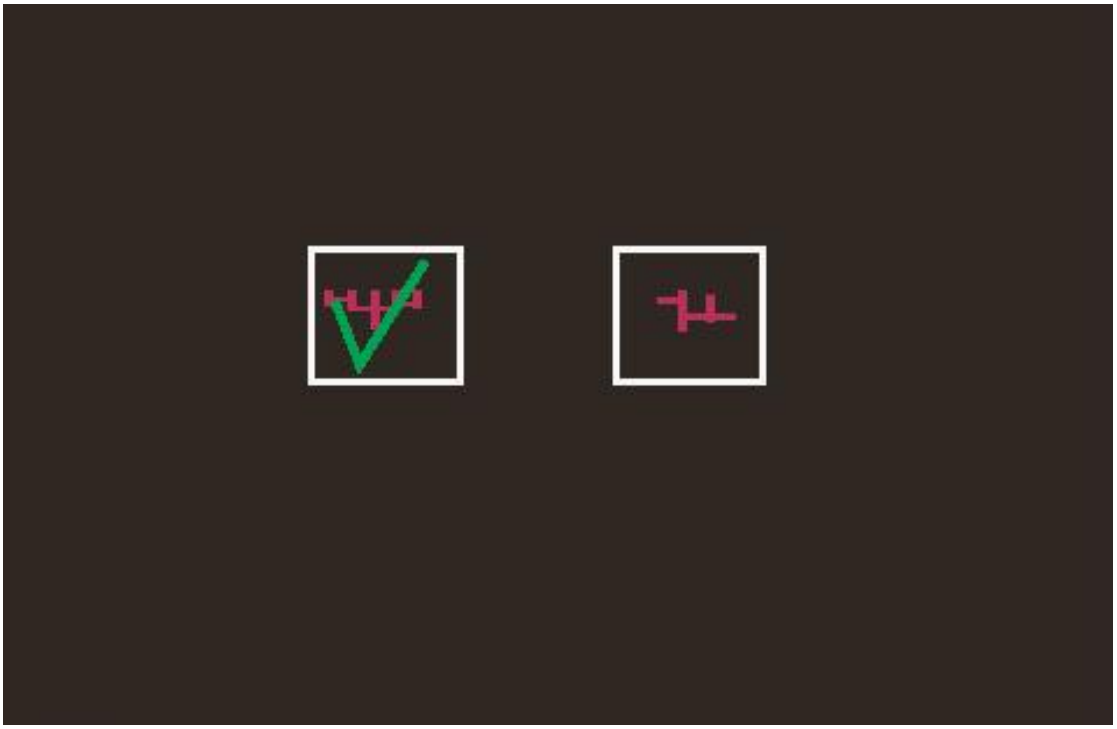


Figure 2.2: CANTAB Spatial Recognition test

2.2.5 Neuroendocrine Testing

In order to profile plasma cortisol and DHEA secretion, subjects were canulated in the antecubital fossa at 12:30 p.m. and blood samples (~5ml) collected at 30 minute intervals from 1:00 p.m. to 4:00 p.m. (Halbreich et al., 1982). Subjects fasted throughout this period, remained semi-supine and did not sleep. Following extraction of serum by centrifugation, samples were immediately frozen and stored at -20°C. This process was carried out by an experienced research nurse.

All assays were performed by staff in the Psychiatry Research Laboratory, Newcastle University, under the supervision of senior technician. Sections 2.2.5.1/ 2.2.5.2 below are descriptions of the assay procedure provided by the senior technician. Details of plasma and saliva methods are presented here (saliva is used subsequently in Chapter 6.2).

2.2.5.1 Cortisol assay

Cortisol was measured using a commercial radioimmunoassay kit (corti-count, ICN). For the determination of cortisol the manufacturer's instructions were followed. Briefly, 25µl of plasma/plasma quality control (QC)/plasma std (27.6 to 1656nmol/l) were incubated at 37°C for 45mins with 500µl of cortisol ^{125}I-tracer in cortisol antibody coated tubes. At the end of the incubation the liquid from each tube was decanted and the tubes inverted and left to drain on absorbent paper for 3mins. Tubes were counted for 1min in a gamma counter (Riastar 5410, Packard). Cortisol concentrations in the samples were determined by interpolation from the standard curve.

QCs (plasma spiked with a known amount of cortisol) were included with each assay and QC rules were applied (Westgard et al., 1981). The mean cortisol values (nmol/l) for plasma (and

saliva) QCs spiked with cortisol were: 102.6 (4.0); 240.7 (21.5); 451.5 (60.8). Intra- and inter-assay CVs (%) for the above plasma (and saliva) QCs were: 5.9/5.6 (10.5%/9.6%); 8.1/8.5 (6.8%/7.0%); 8.1/8.7 (6.2%/8.1%) respectively.

2.2.5.2 DHEA assay

DHEA was measured in extracted saliva/plasma using a modified DHEA tritium radioimmunoassay kit (ICN).

DHEA was extracted from 500µl of saliva into 5ml ethyl-acetate:hexane (3:2 v/v). For plasma, 2ml of ethyl-acetate:hexane (3:2 v/v) was added to 200µl of sample. Extraction QCs (sample spiked with a known quantity of 3H-DHEA) were included to assess extraction efficiency. Typically the recovery from the sample was >80%. The organic phase was removed and evaporated to dryness under a gentle stream of air.

The dry saliva extract was reconstituted in 1.2ml steroid diluent (2.5 ml for plasma extract) and analysed in duplicate as follows. 100µl of 3H-DHEA tracer (ICN) and 100µl of antiserum (ICN) was added to 500µl of reconstituted sample/standard (0-0.5ng/500µl)/QC (steroid diluent spiked with DHEA) and incubated overnight at 4°C. The free ligand was separated from the bound fraction by the addition of 200µl of dextran/charcoal (ICN) which was incubated for 20mins at 4°C before centrifuging at 2500RPM for 15mins at 4°C. The supernatant was decanted into scintillation vials and 3ml of scintillant (Fisher) added before counting for 2mins in a β-counter (Tri-Carb 2100TR, Packard). The cross reactivity of the antibody for DHEA-S was less than 1.2%. DHEA concentrations in the sample were calculated from the standard curve after correction for extraction efficiency. QCs were included with each assay and QC rules applied (Westgard et al. 1981). The calculated values (ng/ml) for QC's used in the DHEA RIA

assay were 0.075, 0.2, 0.5. Actual measured values were 0.06, 0.17, 0.47. The intra- and inter-assay CVs for the entire procedure were 13.0% and 13.5% respectively for saliva; 18.0% and 12.2% for plasma.

2.2.6 Statistical Analysis

2.2.6.1 General data presentation and analysis approach

Descriptive statistics are presented as mean, standard deviation (s.d.) and range. For graphical presentation of results, bar charts are presented as mean, with error bars representing $\pm$ 1 standard error of the mean (SEM). The general approach to data analysis will be parametric. The distribution of the data will be assessed using the Shapiro-Wilk's test and Levene's test for normality and homogeneity of variance respectively. Data will further be visually inspected using histograms and box plots to identify outliers. Where possible, transformations will be applied to data or otherwise supported using the equivalent non-parametric test.

2.2.6.2 Neuropsychological data analyses

A frequent problem associated with studies of neuropsychological function is the number of comparisons that are made in the analysis, increasing the likelihood of committing a type I error. One approach which has been suggested to overcome this is to adopt a multivariate approach to confirm an overall effect of group, prior to examination of individual tests. Similar to all parametric analysis methods there are assumptions about the data underlying the use of this method which include independence, multivariate normality and equality of the population covariance matrices of the dependent variables. However, with respect to Type I error, the MANOVA is generally robust to the latter of these two assumptions (Stevens, 2002).

This method was therefore adopted to confirm that a significant multivariate statistic was present before univariate comparison and presentation of simple effect sizes (Cohen's d). Effects sizes are presented with estimates of the 95%CI for d (Calculator available from the Centre for Evaluation and Monitoring: http://www.cemcentre.org). Further detailed analysis was carried out using analysis of variance (ANOVA) with group (BD or controls) as a between subjects factor. Where a test has more than one level (i.e. Rey-AVLT, learning trials; Vigil, time quarter; SWM, problem level) an additional within subject factor of 'time' or 'level' was added where appropriate. Where ANOVA sphericity was violated according to Mauchly's Test, Huyhn-Feldt epsilon-adjusted significance levels (p_{lhfl}) are reported although unadjusted degrees of freedom are reported for clarity (Field, 2000). Similarly, if homogeneity of variance was violated, the adjusted t and significance levels were reported.

Crawford's method was used to assess if differences qualified as true differential deficits (Crawford et al., 2000). (Also see discussion in section 1.3.1.2).

2.2.6.3 Neuroendocrine data analyses

Data were log (base 10) transformed prior to analysis (Bland & Altman, 1996). Untransformed data summaries are reported for clarity. Cortisol levels, DHEA levels and molar cortisol-DHEA ratios were examined in separate repeated measures ANOVAs with $\log_{10}$ transformed sample (the 7 time points from 1:00 p.m. to 4:00 p.m.) as a within subjects factor and diagnosis (bipolar or control) as a between subjects factor. Where sphericity was violated, within subject degrees of freedom were adjusted using the Huynh-Feldt correction. The adjusted significance values are reported, though the original degrees of freedom are reported for clarity.

Following the primary data analysis, a ROC (Receiver Operating Characteristics) analysis was performed to ascertain the discriminative utility of the endocrine markers i.e. the level at which the greatest separation between the BD and control groups occurred expressed as sensitivity and specificity (Altman et al., 2000).

2.3 Results

2.3.1 Demographic details of the bipolar sample

Twenty bipolar patients (18 male, 2 female) participated in the study. Patients were aged between 26 and 63 years (mean=49 years, s.d.=11) and had no current diagnosis of substance abuse or dependence. There were no current psychotic features in the group. The average (median) age of onset in the group was 20 years (mean=25.5, s.d.=12.5). The median number of hospitalizations in the group was 3. Nine patients (45%) had previously attempted suicide and 7 (35%) had previously been treated with ECT (>12 months ago).

All patients had persistent depressive symptoms, with 17 fulfilling SCID criteria for current depressive episode[4]. The median length of current depressive episode in the group was 7 months (mean=13.5, s.d.=15.7). Depressive symptoms had a mean score of 23 (s.d.=10) on the Montgomery-Åsberg Depression Rating Scale (MADRS) and of 18 (s.d.=10) on the 17-item Hamilton Depression Rating Scale ($HDRS_{17}$). The mean MADRS and $HDRS_{17}$ scores of the three patients without a specific episode were 8 (s.d.=5) and 4 (s.d.=1) respectively. The average YMRS score in the whole group was 4 (s.d.=4).

[4] For brevity in description, the phrase 'bipolar depression' will be used when describing these patients as a group (n=20).

All patients were currently receiving medication at the time of testing which had been stable for at least 6 weeks.

Seventeen patients (85%) were currently taking at least one mood stabilizer. Of the eleven (55%) taking lithium: four were on monotherapy, five were also taking lamotrigine, and two gabapentin (one with valproate, one with carbamazepine). Of the four (20%) taking valproate: three were on monotherapy, one was also taking lamotrigine. Of the two (10%) taking gabapentin: one was on monotherapy and one also taking carbamazepine.

Twelve patients (60%) were currently taking at least one antidepressant: three were taking venlafaxine, three venlafaxine plus mirtazapine, one mirtazapine, one mirtazapine plus citalopram, one amitriptyline, one sertraline, two paroxetine.

Ten patients (50%) were currently taking an antipsychotic: Olanzapine (n=3), Quetiapine (n=4), Risperidone (n=2), Sulpiride (n=1).

2.3.2 Neuropsychological functioning

The matching algorithm, as described in section 2.2.2.2, produced a sample of 20 healthy control participants who were very closely matched to the patient group. Eighteen males and 2 females were selected, matched for age (patients: mean=48.6 years, s.d.=10.8; controls: mean=47.1 years, s.d.=9.3; t=0.488, df=38, p=0.628) and NART score (patients: mean=111, s.d.=6.9; controls: mean=110, s.d.=8.3; t=0.408, df=38, p=0.686). All participants had also been through a screening procedure at the time of testing to exclude anyone with a personal or family history (first-degree) of psychiatric illness, significant medical or neurological illness likely to affect neuropsychological functioning, or history of drug/alcohol abuse.

Initial analysis of the data was performed by MANOVA on the outcome measures from the overall tests battery. To reduce inter-correlation between outcome measures, if multiple possible outcomes were available, then the most commonly utilised measures were included. This resulted in a MANOVA with group (bipolar or control) as the fixed factor and 13 dependent variables (see Table 2-4 for the measures included; note the Immediate Memory tests and Stroop were not included in this due to the reduced control sample available and were analysed separately. Also Rey-AVLT A7 was not included and the percentage A7 retaining measure used). Initially, age and NART were included as covariates although neither were significant and were therefore dropped in the main analysis.

2.3.2.1 Primary analysis

From the primary MANOVA, a main effect of group was observed (Pillai's Trace=0.570; F=2.647, df=13,26, p=0.017) with bipolar patients performing globally worse than controls. The sub-analysis of Immediate Memory and Stroop measures also produced a significant MANOVA main effect of group (Pillai's Trace=0.344; F=3.152, df=4,24, p=0.032). Therefore pairwise analyses were carried out on each of the individual measures (see Table 2-4).

From the pairwise comparisons and examination of effect sizes it is clear that there is a broad impairment across all domains examined, with large effects (d>0.7) in immediate memory (corresponding to visual and spatial WM slave systems Baddeley & Hitch, 1974), executive functioning (spatial working memory, phonological fluency and sustained attention), spatial (but not visual) recognition memory, delayed verbal recall and recognition, and psychomotor speed. Due to the relatively small sample sizes, the distribution of some outcome measures was non-normal, particularly Rey-AVLT recognition and Vigil (in the case of the latter, outliers >3s.d. were also evident in the data). As transforming the data did not alter this, these results

were also examined non-parametrically. The total omission (U=74.0, p<0.001) and commission (U=124.5, p=0.040) errors were significantly greater in the patient group. For the Rey-AVLT recognition trial, the difference between the groups was not statistically significant (U=140.0, p=0.108).

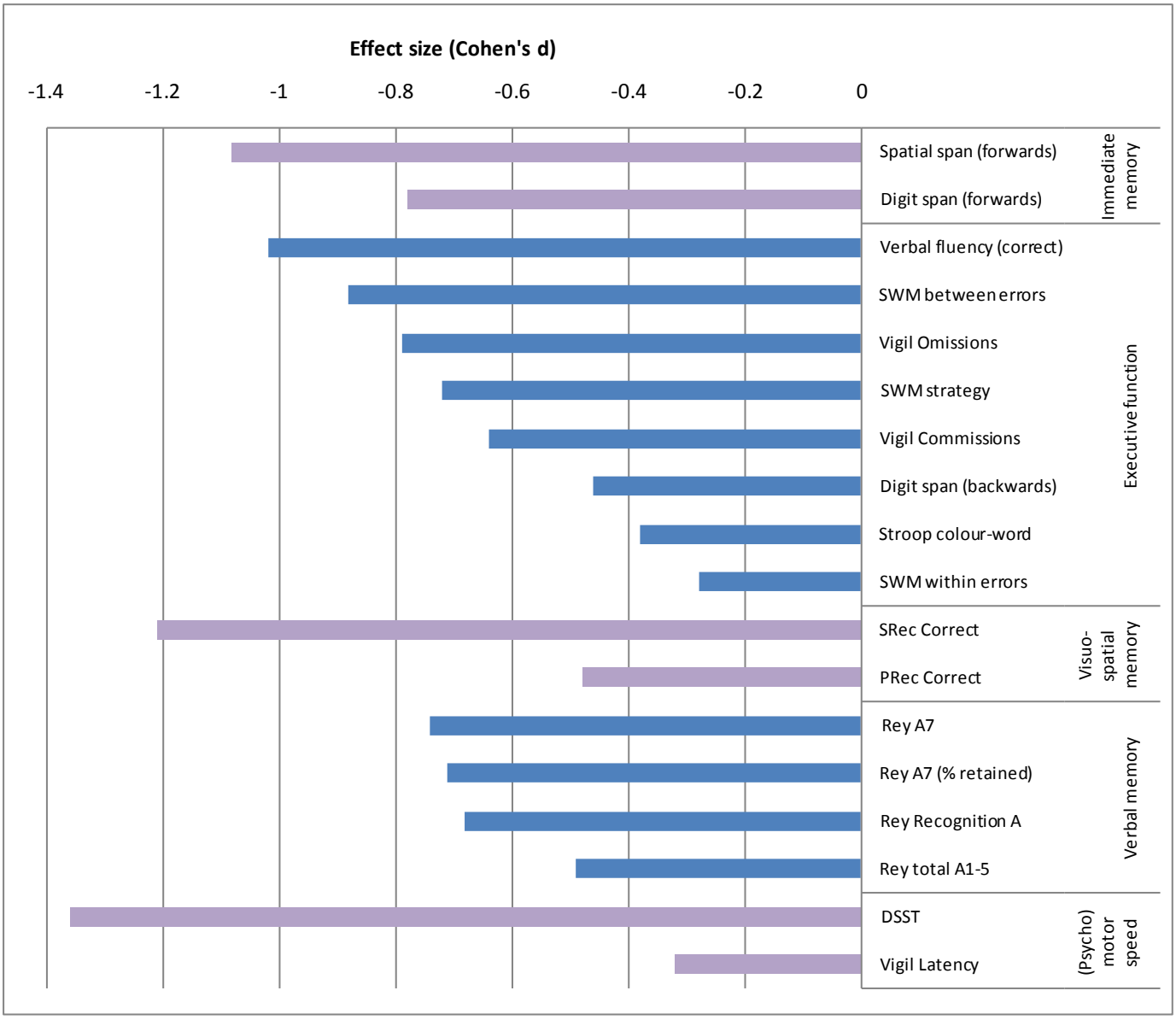
Table 2-4. Neuropsychological data for bipolar patients (n=20) and their matched control group (n=20)

	Bipolar Disorder		Control			95%CI		t	p
	Mean	s.d.	Mean	s.d.	ES				
Immediate memory									
Spatial span † (forwards)	4.9	1.3	6.2	1.0	-1.08	-0.30	-1.82	-3.040	**0.005**
Digit span † (forwards)	6.4	1.1	7.2	0.9	-0.78	-0.02	-1.50	-2.094	**0.045**
Executive function									
SWM between errors	45.1	32.0	23.6	12.8	-0.88	-0.22	-1.51	2.793	**0.010** [a]
SWM within errors	4.6	12.7	2.0	2.6	-0.28	+0.34	-0.90	0.913	0.367
SWM strategy score	35.7	5.3	31.9	5.2	-0.72	-0.07	-1.35	2.290	**0.028**
Verbal fluency (correct)	31.7	12.5	45.0	13.5	-1.02	-0.34	-1.66	-3.217	**0.003**
Digit span † (backwards)	4.4	1.6	5.1	1.4	-0.46	+0.28	-1.17	-1.246	0.223
Vigil Omissions	15.6	23.7	2.2	3.7	-0.79	-0.13	-1.42	2.500	**0.021** [a]
								(U=74.0	***0.001)***
Vigil Commissions	6.4	9.9	1.8	2.3	-0.64	+0.01	-1.26	2.011	0.057 [a]
								(U=124.5	***0.040)***
Stroop colour-word †	85.2	20.9	92.3	15.5	-0.38	+0.38	-1.11	0.996	0.328
Visuo-spatial memory									
SRec memory	13.0	3.8	16.7	2.1	-1.21	-0.51	-1.85	-3.847	**0.001** [a]
PRec memory	19.9	3.7	21.4	2.5	-0.48	+0.16	-1.09	-1.503	0.141
Verbal memory									
Rey total A1-5	42.9	10.9	48.2	10.8	-0.49	+0.15	-1.11	-1.545	0.131
Rey A7	7.3	3.4	9.7	3.1	-0.74	-0.08	-1.36	-2.327	**0.025**
Rey A7 (% retained)	63.7	18.7	76.3	17	-0.71	-0.05	-1.33	-2.231	**0.032**
Rey Recognition A	11.2	3.3	13.0	1.8	-0.68	-0.03	-1.30	-2.130	**0.042** [a]
								(U=140.0	*0.108)*
Psychomotor speed									
DSST	42.9	7.1	57.0	12.8	-1.36	-0.65	-2.02	-4.316	**<0.001** [a]
Vigil Latency	402.1	64.3	383.8	49.5	-0.32	+0.31	-0.94	1.009	0.320

† *Control data available n=12*

[a] *P values reported are adjusted for non-equality of variance following significant Levene's test.*

Figure 2-3. Effect sizes for all neuropsychological test measures



n.b. alternate colouring of bars presented as visual aid only and are to separate neuropsychological domains.

(data corrected so that negative effect sizes always reflect worse performance in the patient sample, relative to controls)

As several measures have multiple 'levels' to their design (i.e. of difficulty or time) these tests were examined in more detail. As the main effect of Spatial Working Memory between search errors was significant, this was examined by repeated measures ANOVA, adding level of difficulty (4, 6 and 8-boxes; see 2.2.4.5.2 for a description) as a within subjects factor. Data were square root transformed prior to analysis.

Significant main effects of group (F=6.083, df=1,38, p=0.018), level (F=200.274, df=2,76, p<0.0001) and a group by level interaction (F=4.663, df=2,76, p=0.012) were observed. Post hoc pairwise comparison across each level revealed that there was no significant difference between the groups at level 4 (p=0.364) with a trend emerging at level 6 (p=0.075) and increasing further to a significant difference at level 8 (p=0.004).

A similar analysis was performed for Vigil omission errors, including time (errors at 2, 4, 6 and 8 minutes) as the within subjects factor. Although the main effect of group was significant there was no significant main effect of time (F=3.305, df=3,114, p_{lhfl}=0.057) or interaction (F=0.977, df=3,114, p_{lhfl}=0.361).

Figure 2-4. Spatial Working Memory between search errors across level of difficulty

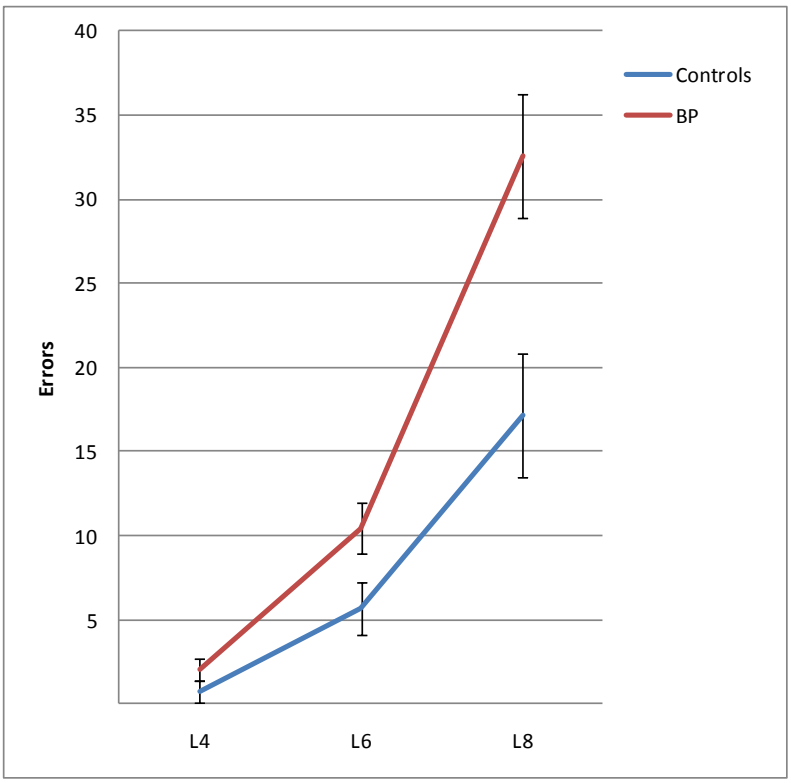
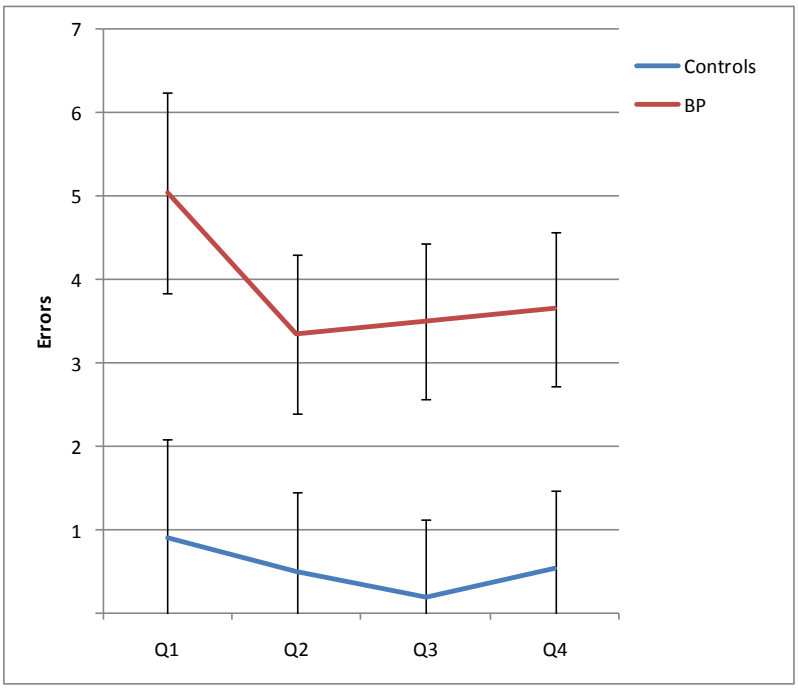
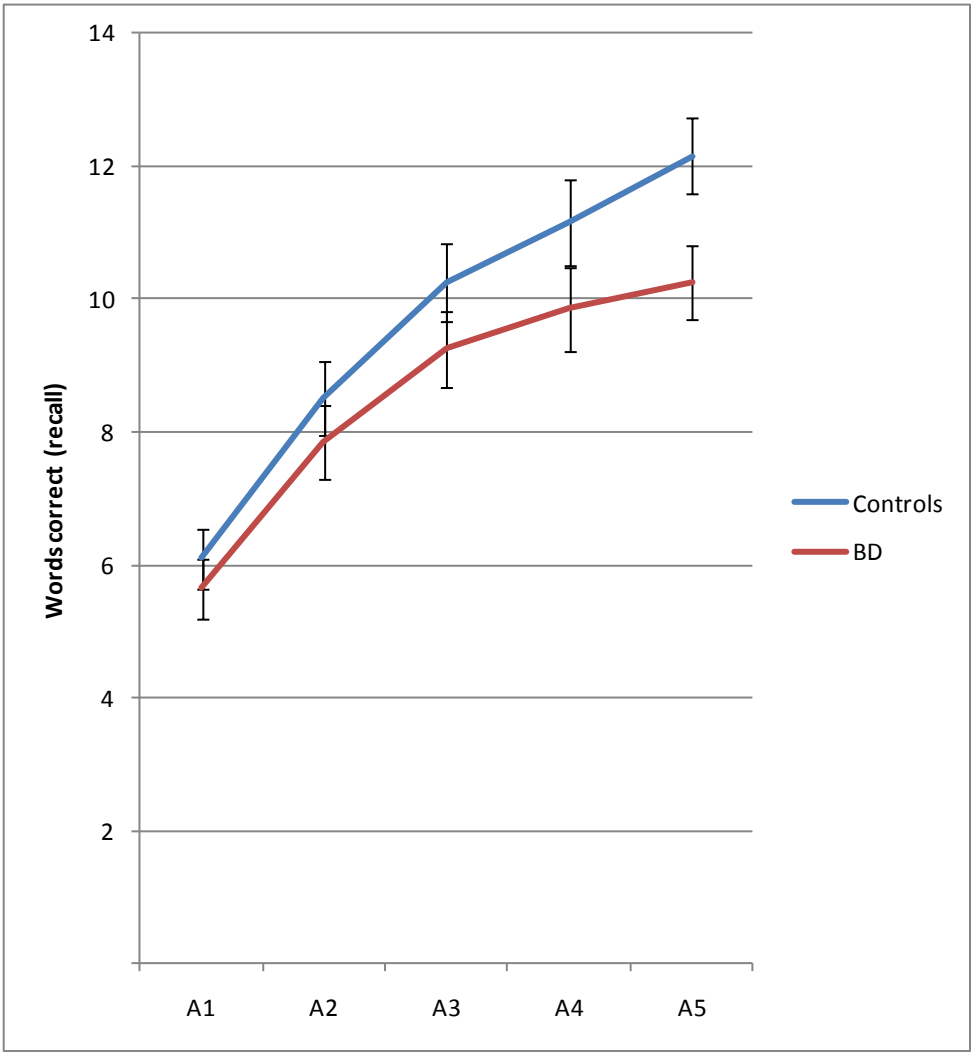


Figure 2-5. Vigil omission errors across each time-quarter of the test



For the Rey-AVLT, a repeated measures ANOVA was also used to examine the 5 repetitions of the list, retaining 'group' as the between subjects factor. The effect of 'list' was significant (F=89.020, df=4,152, $p_{|hf|}$<0.0001) however there was no significant group by list interaction (F=1.612, df=4,152, $p_{|hf|}$=0.187) suggesting that whilst recall improved with increased repetition, this learning/recall was similar across patients and controls.

Figure 2-6. Rey-AVLT recall across each repetition of list A.



2.3.2.2 Exploratory analysis: differential deficit

Examination of the profile of effect sizes revealed some interesting differences that were explored further to see if they represented true differential deficits (i.e. whether the difference or deficit on test A was significantly greater than the deficit on test B).

The first comparison subjected to this analysis was within the immediate memory domain, where a larger effect size was evident in spatial span than digit span. However, this difference did not qualify as a differential deficit (t=0.579, df=29, p=0.284).

Secondly, within the visuo-spatial memory domain, the effect size observed in spatial recognition was larger than that in pattern recognition. This difference did qualify as a differential deficit (t=1.815, df=37, p=0.039)[5].

2.3.3 Neuroendocrine testing

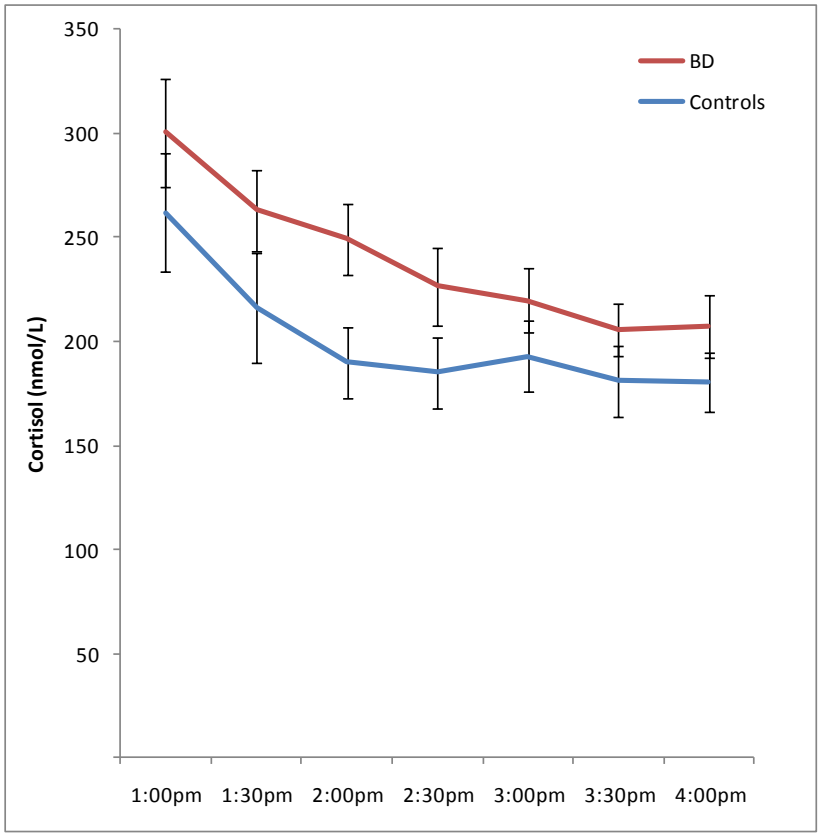
As the control participants in earlier neuropsychological studies had not completed the 1pm to 4pm blood-sampling protocol (and had only completed neuropsychological testing) a sample of 20 healthy controls were recruited from hospital staff and by local advert. All were physically healthy and were subject to the same inclusion/exclusion criteria as in section 2.3.2. The group was matched exactly for sex (18 males, 2 females) and were closely matched for age (mean=45.3 years, s.d.=12.4; t=0.897, df=38, p=0.375) with the patient sample.

The DHEA level at a single time point (3:30 p.m.) was missing for one bipolar patient. This was replaced using the calculated midpoint between the 3:00 p.m. and 4:00 p.m. samples for this subject.

[5] There is a potential issue of ceiling effects in the Pattern Recognition test and therefore this may be an artefact of a reduced potential effect size. A total of 8 participants (5 controls, 3 patients) achieved the maximum score of 24.

The repeated measures ANOVA on cortisol levels revealed a significant main effect of group (F=4.339, df=1,38, p=0.044) with patient having higher levels than controls. There was clear evidence of diurnal rhythm in a significant main effect of time (F=8.875, df=6,228, p<0.0001) with levels decreasing over the afternoon, although there was no group by time interaction (F=0.517, df=6,228, p=0.678).

Figure 2-7. Plasma cortisol levels from 1pm to 4pm in patients and controls



For DHEA, no significant main effects were observed for group (F=0.511, df=1,38, p=0.479), time (F=2.043, df=6,228, p=0.084) or in the group by time interaction (F=0.094, df=6,228, p=0.988). Similarly, for the molar cortisol-DHEA ratio analysis, there were no significant main effects of group (F=0.574, df=1,38, p=0.453) or group by time interaction (F=1.086, df=6,228, p=0.356) although there was a significant diurnal rhythm, reflected in the main effect of time (F=11.938, df=6,228, p<0.0001).

Figure 2-8. Plasma DHEA levels from 1pm to 4pm in patients and controls

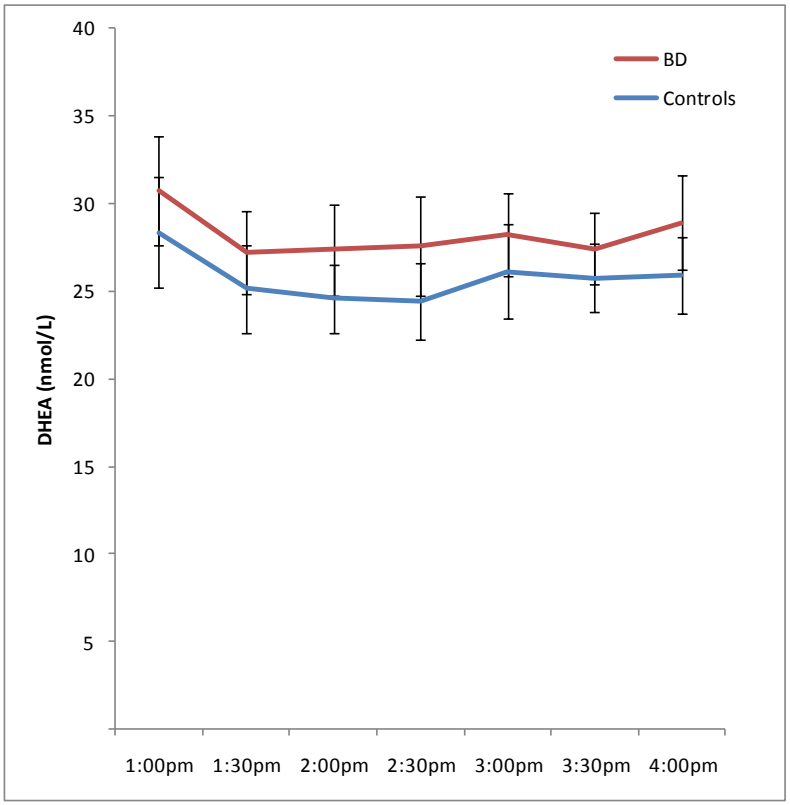
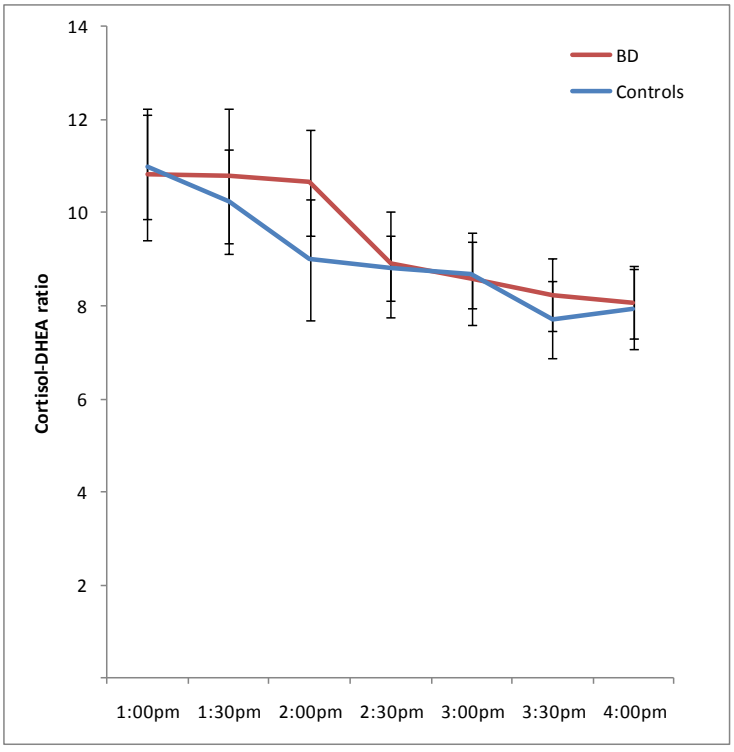


Figure 2-9. Plasma Cortisol-DHEA ratios from 1pm to 4pm in patients and controls



2.3.3.1 ROC analysis

The ROC analysis for cortisol, DHEA and cortisol-DHEA ratio is presented in the table below. Only cortisol levels had a 95%CI above 0.5 (random classification) and using a cortisol AUC level of 33,323.16 nmol/L/min gave good sensitivity with reasonable specificity. Although DHEA and the cortisol-DHEA ratio also gave high sensitivity, the specificity was poor and neither produced AUC values that indicated they were reliable discriminators of the patient and control groups.

Table 2-5. ROC analysis for endocrine data

	Cortisol	DHEA	Cortisol/DHEA ratio
Bipolar *vs*. Control group	AUC=0.72 (95%CI=0.53 to 0.90) Sensitivity=0.90, Specificity=0.60	AUC=0.55 (95%CI=0.35 to 0.80) Sensitivity=0.95, Specificity=0.20	AUC=0.58 (95%CI=0.38 to 0.78) Sensitivity=0.90, Specificity=0.45

2.3.3.2 Exploratory correlations

To explore the relationship between endocrine measures and neuropsychological tests, Spearman's correlations were performed. Only 2 of the 18 correlations performed were statistically significant using cortisol AUC and 1 of 18 using DHEA. None of these would remain significant with correction for multiple comparison with correlation coefficients (see Curtin & Schulz, 1998).

Table 2-6. Spearman's correlation coefficients for endocrine measures and neuropsychological tests in bipolar patients (n=20)

	Cortisol AUC	DHEA AUC
Spatial span	**0.549***	-0.151
Forward span	-0.100	-0.249
SWM between error	-0.141	0.309
SWM within error	0.255	0.113
SWM strategy score	-0.201	-0.064
FAS correct	-0.313	-0.081
Backward span	**0.611****	0.170
Vigil Omissions	0.103	-0.070
Vigil Commissions	-0.233	-0.239
Stroop CW correct in 2min	-0.074	-0.404
SRec Correct	-0.027	-0.194
PRec Correct	-0.089	-0.336
Rey total A1-5	-0.036	-0.324
Rey A7	0.056	-0.323
Rey A7 % retained	0.179	-0.191
Rey Recognition A	-0.196	**-0.634****
DSST	-0.095	-0.253
Vigil Latency	-0.006	-0.310

* p<0.05, **p<0.005 (uncorrected for multiple comparison)

2.4 Discussion of Chapter 2

In this chapter, data were presented from the assessment of neuropsychological and neuroendocrine functioning in 20 patients with bipolar depression and matched groups of healthy controls. From the neuropsychological assessment, there was evidence of broad impairment across all domains examined, with large effects in immediate memory (phonological and spatial), executive functioning (spatial working memory, phonological fluency and sustained attention), spatial (but not visual) recognition memory, delayed verbal recall and recognition, and psychomotor speed. With regard to the neuroendocrine assessment, significantly elevated cortisol levels were evident in patients compared to the control group, although there was no significant difference in DHEA or cortisol-DHEA ratio. Cortisol levels (AUC; 1pm to 4pm) were also showed moderate discriminatory value, yielding sensitivity of 0.90 and specificity of 0.60 in a ROC analysis. There was no evidence of a simple linear or monotonic relationship between neuropsychological and neuroendocrine parameters, with an absence of any meaningful statistically significant correlations.

Comparing the profile of mean effect sizes obtained in this study with those of the pooled estimates of effects sizes from the literature in Chapter 1.3.2, several fall within the 95%CI including digit span forwards, Spatial Working Memory, and delayed verbal recognition. Tests that produced values lower than the pooled estimates were reverse digit span, Stroop (inhibition), Rey-AVLT immediate and delayed recall ; while those that produced higher values were Spatial Span, Spatial Working Memory, Spatial and Pattern Recognition and DSST. However, this does not account for a similar confidence parameter applied to the present data, which would result in some degree of overlap in all measures (see Figure 1-2; section 1.3.8). One very important consideration is that we do not know the extent of biological dysfunction, specifically HPA axis function, in most studies. If HPA dysfunction is causal to neuropsychological impairment (perhaps, though, not monotonically as we have assessed

here), this factor may at least in part explain the discrepancy between the present findings and previous literature.

It is of note that the majority of tests that produced effect sizes larger than those from the literature review involved spatial processing. As discussed in the introduction, several studies have indicated that certain clinical feature can exacerbate the profile and extent of neuropsychological impairment. For example, Glahn and colleagues (2006) reported that bipolar depressed patients with history of psychosis exhibited spatial memory deficits while those without did not. Unfortunately data is not available on history of psychosis in these patients in the present thesis, although those recruited for the study were through secondary care and a specialist tertiary referral service and therefore are likely to have more complex, chronic illnesses. However, this is unlikely to be the only reason for the profile of results obtained as it does not explain the occurrence of effect sizes lower than those obtained from the pooled literature search.

With regard to the cortisol data, despite the high sensitivity and moderate specificity in separating the patient and control groups, levels did not correlate to any great extent with neuropsychological outcome measures. This highlights a potential difficulty in establishing a simple linear or monotonic relationship and the need for careful consideration of firstly, the known effects of corticosteroids on specific rather than general cognitive processes and secondly, the subtleties of HPA axis dysfunction. Given the potential complexity of these relationships, it is not surprising that it is difficult to establish a simple (and replicable) linear or monotonic model relating peripheral cortisol levels to broad neuropsychological functions. It is possible that the simple assessment of basal levels is not representative of the dynamic processes or the role of individual receptors in cognitive processes. Perhaps the linking of the 'activated' HPA axis is a better method and may relate to neuropsychological functioning in a

more representative way. This includes the function of the GR specifically which plays an important role in memory (see section 1.4). Therefore assessment after activation or blockade of the GR may be a plausible method.

The next chapter of this thesis therefore reports the results of a study in the same patients in which a GR antagonist or placebo was administered for one week and neuropsychological and neuroendocrine functioning was assessed. It was hypothesised that administration of the drug would acutely raise cortisol levels but subsequently result in a reduction once treatment had ceased and that this change would selectively improve neuropsychological functioning.

Chapter III

Effects of antiglucocorticoid treatment on neuropsychological functioning, mood and the HPA axis in bipolar depression

3. Effects of antiglucocorticoid treatment on neuropsychological functioning, mood and the HPA axis in bipolar depression

3.1 Introduction to Chapter 3

Although the co-occurrence of HPA axis dysfunction and neurocognitive impairment has frequently been described (see chapter 1.5), demonstrating a direct causal link has proved more difficult. As noted above, there are a number of possible reasons for this, not least of which may be that the complexity of the relationship centrally between the two phenomena may not be accurately modelled or measured by peripheral hypercortisolaemia. In this section studies that have used more direct probes of glucocorticoid receptor function to assess effects on learning and memory are examined.

One opportunity lies in the novel use of GR antagonist drugs which are currently being explored for the treatment of severe mood disorders. The intention here is not to review this literature but to very briefly discuss this approach with relevance to the application in neuropsychological functioning in mood disorder, focussing specifically on one drug of interest – mifepristone.

3.1.1 Mifepristone - background

Mifepristone (or RU-486) is a synthetic steroid with both antiprogesterone and antiglucocorticoid properties. The compound is a 19-nor steroid with substitutions at positions C11 and C17 $[17\ \text{beta} - \text{hydroxy} - 11\ \text{beta} - (4 - \text{dimethylamino phenyl})\ 17\ \text{alpha} - (1 - \text{propynyl})\text{estra} - 4{,}9 - \text{dien} - 3 - \text{one}]$ which antagonizes cortisol action competitively at the receptor level (Nieman et al., 1985). It was discovered in the early 1980s by the French pharmaceutical company Roussel–Uclaf (Herrmann et al., 1982; Jung-Testas & Baulieu, 1983).

At present it is licensed in the UK for the medical termination of pregnancy (trade name: Mifegyne®; marketing authorization holder: Exelgyn Laboratories). Mifepristone was the first antiprogestin to be developed and it has been evaluated extensively for its use as an abortifacient. The original target for the research group, however, was the discovery and development of compounds with antiglucocorticoid properties (Hazra & Pore, 2001), and it is these properties that are of greatest interest for their application in the treatment of severe mood disorders and psychosis.

3.1.2 Pharmacokinetics and pharmacodynamic activity

The pharmacokinetics of mifepristone are dose-dependent in humans (Ashok et al., 2002). Due to saturation of the serum-binding capacity, high dose mifepristone results in nonlinear kinetics, whereas lower doses show a linear pattern (Leminen et al., 2003). For example, following administration of doses from 50 to 800mg, after the absorption and distribution phase of approximately 4 to 6h, the serum concentration of mifepristone remains in the micromolar range for the next 24 to 48h. Within the dose range of 2 to 25mg, serum concentrations of mifepristone, as well as the areas under the concentration–time curves (AUC), increase according to dose (Sitruk-Ware & Spitz, 2003).

Following a single oral dose of 600mg mifepristone, the binding equivalent is present in measurable concentrations 7 days after administration, only decreasing below assay detection limits >7 to 14 days (Foldesi et al., 1996). In this study, the concentration of the mifepristone binding equivalent reached a peak within approximately 2 hours (doses 200 to 600mg) indicating rapid absorption. Peak levels were significantly greater following the 600mg dose (Cmax=12.3µmol/L vs. 200mg: 6.30µmol/L), while the bioavailability as assessed by the AUC was significantly greater following 600mg dose than both 200 and 400mg. These were not, however, directly proportional to the dose increase (Foldesi et al., 1996).

In contrast to mifepristone plasma concentrations, plasma concentrations of its metabolites do increase in a dose-dependent manner when larger doses are administered, so that serum metabolite concentrations being close to, or even in excess of those of the parent compound (Lahteenmaki et al., 1987). These metabolites have some antiprogestin and antiglucocorticoid properties, and therefore may mediate some of the actions of mifepristone (Spitz & Bardin, 1993a, 1993b).

3.1.3 Side effects of chronic mifepristone administration

Laue and colleagues reported that in healthy male normal volunteers who received mifepristone (10 mg/kg/day), 8 of 11 subjects developed generalized exanthem after 9 days. One subject developed symptoms and signs consistent with the diagnosis of adrenal insufficiency (Laue et al., 1990). With respect to immune function, it was reported that total white blood cell counts, absolute lymphocyte, neutrophil and eosinophil counts, erythrocyte sedimentation rate, and quantitative immunoglobulins did not change. Furthermore, functional evaluation of lymphocyte cytotoxicity and proliferation revealed no changes.

A study using lower doses (200mg/day for 2 to >31months) in 14 patients with unresectable meningiomas reported milder side effects. Most commonly, fatigue was noted in 11 of the 14 patients (Grunberg et al., 1991). However, in a study of mifepristone (200mg/day for up to 8 weeks) in chronic depression, 1 of 4 patients discontinued treatment prematurely because of the appearance a rash (Murphy et al., 1993). In patients with psychotic depression receiving mifepristone (50 to 1200mg/day for 7 days), 2 of 10 patients in the 600-mg group and 1 of 9 in the 1200-mg group reported uterine cramping, while 1 of 11 patients in the 50-mg group and 1 of 9 patients in the 1200-mg group (but none in the 600-mg group) reported a rash. In both cases, this had abated 1 to 2 months after study completion (Belanoff et al., 2002).

3.1.4 Antiglucocorticoid effects of mifepristone

A large amount of human clinical data on the antiglucocorticoid actions of mifepristone has come from studies in Cushing's disease (Sartor & Cutler, 1996). Nieman and colleagues administered mifepristone orally at increasing doses of 5, 10, 15, and 20 mg/kg/day for a 9-week period to a patient with Cushing's syndrome due to ectopic ACTH secretion. Following treatment, the somatic features associated with Cushing's syndrome ameliorated and blood pressure normalized. Importantly, suicidal ideation and depression also resolved, and all biochemical glucocorticoid-sensitive parameters normalized (Nieman et al., 1985).

Mifepristone has also been shown to rapidly reverse acute psychosis in Cushing syndrome (van der Lely et al., 1991). More recently, high-dose (up to 25 mg/kg/day), long-term mifepristone administration was shown to normalize all biochemical glucocorticoid-sensitive measurements, as well as significantly reverse psychotic depression in a patient with Cushing's syndrome caused by an ACTH-secreting pituitary macroadenoma (Chu et al., 2001). Although the adrenal axis also normalised, the 18 month-long mifepristone treatment course led to the development of severe hypokalemia (attributed to excessive cortisol activation of MRs), which responded to spironolactone administration.

3.1.5 The use of mifepristone in mood disorders

Early work highlighted the potential for antiglucocorticoid strategies in depression. Initially the focus of studies utilising mifepristone was on the effect on endocrine parameters (Kling et al., 1989; Krishnan et al., 1992). In the first preliminary, open-label investigation of mifepristone treatment of major depression, Murphy and colleagues administered mifepristone (200mg each morning) for as long as it was tolerated, for up to 8 weeks to 4 patients with 'drug-resistant' depression. Data were presented as a case-series and showed improvements of

between 16% and 66% on the Hamilton Depression Rating Scale (HDRS) (Murphy et al., 1993). The trial terminated, however, due to problems obtaining the trial medication.

Recent studies have renewed interest in the potential therapeutic efficacy of GR-antagonists in the treatment of mood disorders and psychosis (for a review see Gallagher et al., 2008). There is considerable debate in the literature currently as to the true efficacy of this approach and some authors have questioned the methodology and analysis of these trials (Carroll & Rubin, 2006, 2008). However, this is outside the scope of specific interest for the present thesis and will not be discussed here. The key point is that mifepristone is a potent modulator of the HPA axis and GR, although no studies have examined the effect of this on neuropsychological functioning in patients with mood disorders. Evidence does however exist on the effects of GR manipulation on cognitive functioning in other groups.

3.1.6 Effects of GR manipulation on cognitive function

3.1.6.1 Effects of GR agonists on memory

Numerous studies have examined the effects of GR agonists on cognitive function, both in animal and human studies. Generally, two approaches have been adopted: the first is to use the cortisol response to specific GR agonists, such as dexamethasone, and the relationship of post-administration cortisol levels with memory; the second is to assess the direct effects of GR agonists/antagonists within the stages of information processing and memory formation.

The use of dexamethasone (and the dexamethasone suppression test; DST) as an assay of GR function has a long history in the assessment of patients with mood disorders (Carroll, 1982b, 1982a; Ribeiro et al., 1993; Nelson & Davis, 1997; Raison & Miller, 2003). Relating this to memory function (Wauthy et al., 1991) has produced mixed results. In elderly depressed

patients, contrary to expectations a failure to dexamethasone-suppress was associated with better global cognitive performance (Adler & Jajcevic, 2001). In contrast, a study from our group found, in bipolar disorder, dexamethasone non-suppression was correlated with poorer working memory function (Watson et al., 2006b). Interestingly, in a series of three studies using the same verbal memory task, Wolkowitz and colleagues found performance deficits in healthy subjects following 5-days of prednisolone (a non-specific MR and GR agonist) administration, in healthy subjects following a single dose of dexamethasone, and in patients with major depression with dexamethasone non-suppression (Wolkowitz et al., 1990). Together these findings are consistent with the inverted "U"-shape relationship between serum glucocorticoid levels and cognitive function posited earlier. This was shown more clearly recently in patients with major depression using a verbal declarative memory task (paragraph recall) administered twice, with dexamethasone or placebo administered in-between (Bremner et al., 2004). In healthy controls, memory improved from baseline to day 3 with placebo but was unchanged with dexamethasone, whereas in MDD patients memory function showed a pattern of decreasing with placebo and improving with dexamethasone (Bremner et al., 2004).

Studies in clinical populations using specific MR and GR probes would be of great interest in examining the hypothesis of MR/GR balance and optimal memory performance (Tytherleigh et al., 2004).

3.1.6.2 Effects of GR antagonists on memory

The recent therapeutic interest in anti-glucocorticoids for the treatment of mood disorders has provided a further valuable opportunity to examine the effects of GR manipulation on neuropsychological functioning (Gallagher et al., 2008). In this final section we will provide an

overview of the findings in the animal literature before examining the parallels in clinical studies in humans.

3.1.6.2.1 Rodent studies

Douma and colleagues examined the effects of repeated MR- (RU28318), GR- (RU38486), or combined antagonism on aspects of spatial learning in the rat. Repeated administration of the MR-antagonist impaired reference memory (in the hole-board learning paradigm). Combined MR/GR-antagonism similarly reduced reference memory performance while GR-blockade alone had no effect. These results highlight the importance of MRs in this process. Working memory acquisition rates were also suppressed in the initial phase of the training period with MR-blockade, although they were impaired throughout the whole training period with combined MR/GR-blockade suggesting modulation by both (Douma et al., 1998). Such results highlight the importance of considering the specific processes within a cognitive domain when examining the relative contribution of receptor function.

As with the human literature discussed above, the pattern of HPA interventions and the timing of assessment of memory processes are crucial. In a series of studies, Oitzl and colleagues have examined the effect of GR blockade on spatial memory function in the rat. Acute intracerebroventricular (i.c.v.) injection of the GR antagonist RU38486 was found to result in spatial memory impairment (Oitzl & de Kloet, 1992). However more localized administration (10 and 100 ng) intra hippocampally was found to improve performance in a water maze 24 h after treatment. This effect occurred following either unilateral or bilateral injection and appeared to be dose-related (Oitzl et al., 1998a). Interestingly the authors note that opposite effects on neuroendocrine regulation of pituitary ACTH release occur with RU38486. While i.c.v. administration increases plasma ACTH and corticosterone levels,

administration locally in the dorsal hippocampus suppresses the circadian rise of these hormones (van Haarst et al., 1997). Chronic administration also has effects that are dependent on the administration regimen. Phasic GR blockade (RU38486: 10 and 100ng/µL i.c.v. administered pre-training over 3 consecutive days) impaired spatial memory in a dose-dependent fashion while continuous blockade (10 and 100ng/0.5µL per hour over 10 days) facilitated spatial performance, continuing several days after training in the case of the higher dose (Oitzl et al., 1998b).

The use of such compounds in recent clinical trials offers an important avenue of research into the role of the GR in human memory. Of particular interest is in determining if effects seen in the non-human literature can be replicated.

3.1.6.2.2 Human studies

There is little information of the effects of GR antagonists on memory function in humans. Most data comes from treatment studies using RU38486 (mifepristone). Pomara and colleagues examined the use of mifepristone in the treatment of Alzheimer's disease in a small randomized, placebo-controlled trial (Pomara et al., 2002). Some subtests of cognition were improved following RU486 on an intention-to-treat analysis. At 12 hours after the first dose, the change from baseline in the Alzheimer's Disease Assessment Scale–Cognitive subscale (ADAS-Cog)(Rosen et al., 1984) total score was not statistically different between the two groups. However, patients treated with mifepristone performed significantly better on the ADAS-Cog Word Recall subtest. At week 6, the mean change from baseline in ADAS-Cog total score among completers revealed that patients treated with mifepristone tended to improve (by 2.67 points) whereas patients treated with placebo tended to worsen (by 1.67 points).

It is important to consider the treatment regimen used in these trials. Although the administration of mifepristone was once daily, more chronic receptor occupancy was likely by the end of the treatment period because mifepristone, has rapid absorption, a long half-life (of 25 to 30 hours) and micromolar serum concentrations following typical doses (Heikinheimo et al., 2003a). This may explain the positive effects found in spatial memory processes, akin to those found in chronic (but not phasic) administration in rodents (see Oitzl et al., 1998).

3.1.7 Summary and aims of the study

From the literature reviewed, from healthy volunteer work through to clinical conditions and studies examining the direct effects of GR antagonists on memory in animals and humans, a common finding is for spatial memory functions and verbal declarative memory to be consistently altered. As was observed in the patient sample participating in the study reported in Chapter 2, large effects were observed for neuropsychological tasks within the spatial domain with some even reaching criteria to be considered a differential deficit (see 2.3.2.2). Given the extent of the HPA axis disturbance in the patient group, particularly in cortisol levels, then we can at this stage hypothesize that giving a drug that targets the GR and potentially reduces hypercortisolaemia will lead to improvements in spatial memory and verbal declarative memory. Therefore, the a priori selection of primary outcome measures were the spatial working memory test from CANTAB and verbal declarative memory, using the Rey-AVLT. The same neuropsychological test battery as in Chapter 2 was used. It is hypothesised that administration of the GR antagonist mifepristone would result in improved neuropsychological functioning.

3.2 Methods

3.2.1 Subjects

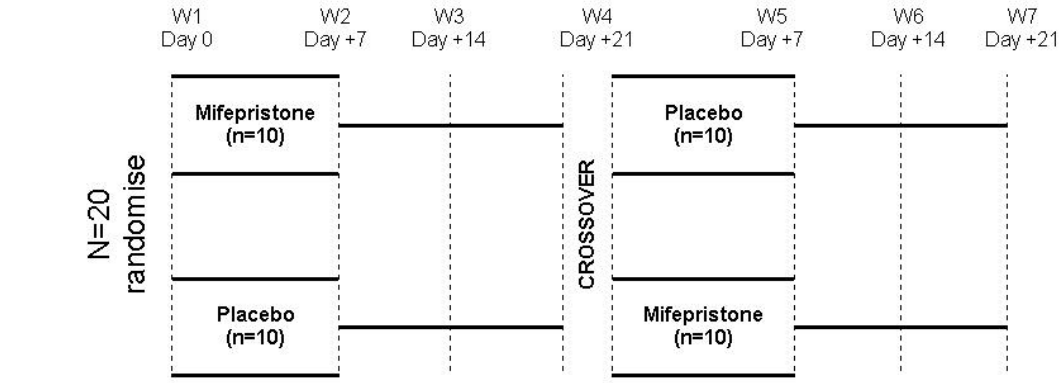
The details of the 20 patients participating have been presented earlier (Chapter 2.2). Patients were aged 18 to 65 years with a diagnosis of bipolar disorder, confirmed using the Structured Clinical Interview for DSM-IV (SCID) (First et al., 1995), were recruited from services in North East of England. A specific attempt was made to recruit those with residual depressive symptoms. Illness characteristics, clinical ratings and medication history were determined by trained psychiatrists using full history, case-note and medication review and standardized rating scales. Patients' medication had been unchanged for 6 weeks prior to participation and remained so throughout the study period. Seventeen were taking at least one mood stabilizer, with 13 taking at least one antidepressant and 11 taking an antipsychotic.

After a complete description of the study, written informed consent was obtained from all participants; the study received full approval from the local ethics committee.

3.2.2 Procedure

Following an initial baseline assessment of neurocognitive function and mood, and basal neuroendocrine profiling (day 0), patients were randomly allocated to receive either 600mg mifepristone (taken orally at 8:00 a.m. once a day) or placebo for 7 days. Administration of medication was in a double-blind design. Mood ratings were taken after the week's treatment (day +7) and then at weekly intervals (day +14 and day +21). At day 21, the groups crossed over and the alternative treatment (placebo or mifepristone) administered for 7 days, again with ratings taken following the week's treatment (day +7) and at weekly intervals (day +14 and day +21). Neurocognitive function was assessed on three occasions over the study period: at baseline and at day +21, after each treatment. Neuroendocrine profiling was performed at baseline, after the week's treatment period (day +7) and then day +21.

Figure 3-1. Trial design



Neuropsych:	Y			Y			Y
Bloods:	Y		Y		Y		Y
Symptoms:	Y		Y	Y	Y	Y	Y
	Baseline						

3.2.3 Neurocognitive testing

Based on previous research on the effects of corticosteroids on neurocognitive function (Lupien et al., 1999; Newcomer et al., 1999; Young et al., 1999; de Quervain et al., 2000; de Quervain et al., 2003), it was predicted that the principal cognitive domains which would be most sensitive to changes in HPA axis function were (spatial) working memory and verbal declarative memory. The primary neurocognitive battery therefore consisted of two tests:

The Spatial Working Memory task. This computerized test of working memory from the Cambridge Neuropsychological Test Automated Battery (CANTAB; CeNeS Pharmaceuticals, Cambridge, U.K.) requires subjects to search through an increasing number of (three, four, six, and eight) boxes to locate hidden tokens. As the token is never located in the same box more than once, "between search errors" (BSE) are committed when the subject returns to search a box in which a token has previously been located.

The Rey-Auditory Verbal Learning Test (Rey-AVLT). This test of verbal learning, includes indices of initial and delayed recall and recognition. A list of 15 words (List A) is read out to the subject 5 times. They are required to recall after each trial. A different list of 15 words (List B) is then read once, followed by recall of this list. Finally, after a 30 minute delay, recall of List A is again tested without an additional presentation of that list. This is followed by a recognition trial of words from List A. The number of words correctly recalled or recognized is recorded. Alternative forms of the test were used on each visit.

A secondary battery was also included which examined a broader range of neurocognitive domains, incorporating additional measures of learning and memory, attention and executive function:

Short-term memory span. This was tested across both phonological and spatial domains. The Wechsler forward digit span test requires subjects to repeat verbatim a string of digits which sequentially increases in length until the consecutive failure of two trials of the same digit span length. The CANTAB spatial span task was utilized to assess the subjects' ability to remember a serial sequence of squares as they change colour.

Visuo-spatial learning and memory. This was assessed using the CANTAB pattern and spatial recognition tests. The pattern recognition task requires the subject to learn a series of 12 abstract patterns. Subjects are then presented with pairs of patterns and required to identify the familiar one. The test consists of two sets of 12 stimuli. For the spatial recognition test, the subject must learn the on-screen spatial position of 5 serially presented squares, with a subsequent forced-choice recognition between 2 locations. A total of four trials of five stimuli are completed. Alternative forms of both tests were used on each visit.

Executive function. This was tested using an established verbal fluency test (naming words beginning with one of three given letters; 60 seconds for each) with the overall total correct responses recorded. The Wechsler backward digit span, which requires the monitoring of information held in working memory, was also administered using the same method as the forward span test. Alternative forms of both tests were used on each visit.

Attention/psychomotor speed. Psychomotor speed was assessed using the digit symbol subtest from the Wechsler Adult Intelligence Scale; a test requiring rapid copying of symbols paired with numbers in 90 seconds. Alternative forms of the test were used on each visit. Aspects of attention were measured using a computerized continuous performance task – Vigil (Cegalis & Bowlin, 1991). In this random-interval 'A-K' form, subjects are required to respond to the target letter 'K' only when it is preceded by the letter 'A' from amongst a stream of random letters over an 8 minute period.

All pen-and-paper tasks were administered according to standardized instructions referenced under each test above and computerized tests from the CANTAB according to the manual protocols, on a personal computer fitted with a colour touch-screen monitor. For all subjects, testing began at 1:00 p.m. and took approximately 75 minutes to complete.

3.2.4 Symptoms

With respect to symptomatic improvement, the antidepressant effect of mifepristone was the principal focus, therefore the outcome measures of interest were the 17-item Hamilton Depression Rating Scale ($HDRS_{17}$; (Hamilton, 1960) and the Montgomery-Åsberg Depression Rating Scale (MADRS; (Montgomery & Asberg, 1979). Other secondary scales consisted of the Brief Psychiatric Rating Scale (BPRS; (Overall & Gorham, 1962) and the Young Mania Rating Scale (YMRS; (Young et al., 1978).

3.2.5 Neuroendocrine assessment

To profile plasma cortisol secretion, subjects were canulated in the antecubital fossa at 12:30 p.m. and blood samples collected at 30 minute intervals from 1:00 p.m. to 4:00 p.m. Subjects fasted throughout this period, remained semi-supine and did not sleep. Cortisol levels were determined by using Corti-cote radioimmunoassay kits (ICN Pharmaceuticals, Costa Mesa, Calif.). The inter-assay coefficient of variation for cortisol was less than 8%, and the intra-assay variation was less than 9% across the assay range.

3.2.6 Statistical analysis

Neurocognitive data were analyzed by repeated measures analysis of covariance (ANCOVA) with 'treatment' (mifepristone or placebo) and, where tests had more than one level, 'level', as the within subject factors. As differential learning effects may occur depending upon the order of treatment administration, 'order' (mifepristone first or placebo first) was entered as a between subjects factor and 'baseline' performance as a covariate. Main effects were further examined as the mean difference (and 95% confidence interval of the difference) between treatments (mifepristone or placebo), expressed as a change from baseline performance (Altman et al., 2000). Mood symptoms were also expressed as the mean change (95%CI) from baseline for each treatment and analyzed by paired t-test. All cited p values were two-tailed, with a significance level set at 0.05. Analyses were performed using SPSS (SPSS, 1998).

3.3 Results

One patient was excluded from the study because of self-discontinuation of lithium prophylaxis. Data from 19 patients were available for analysis.

3.3.1 Neuropsychological testing

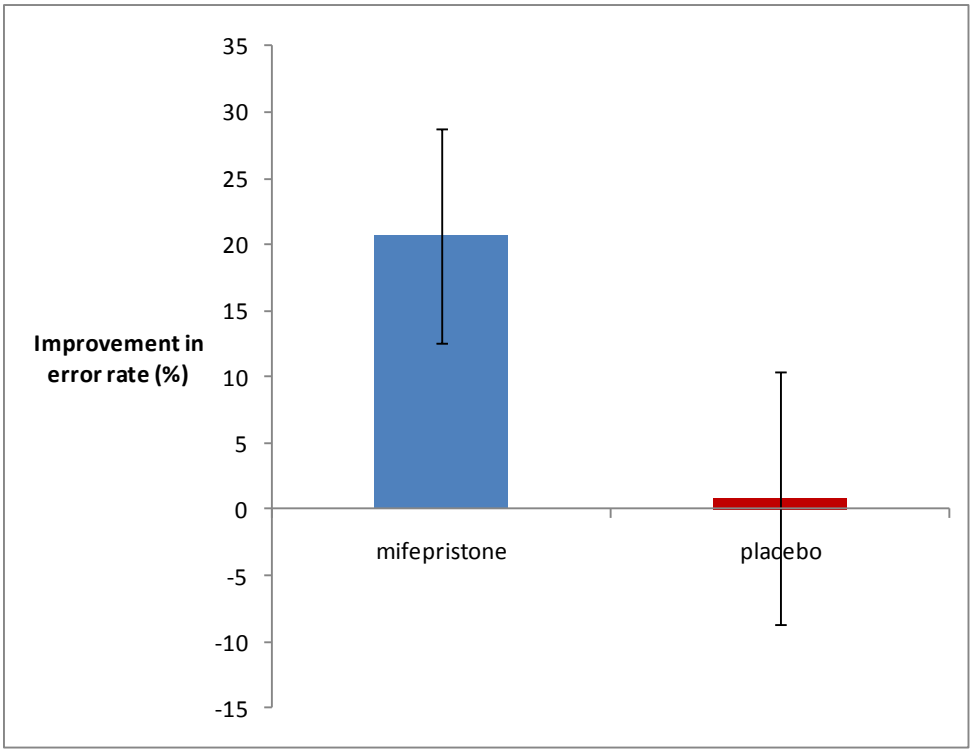
3.3.1.1 Primary outcome measures

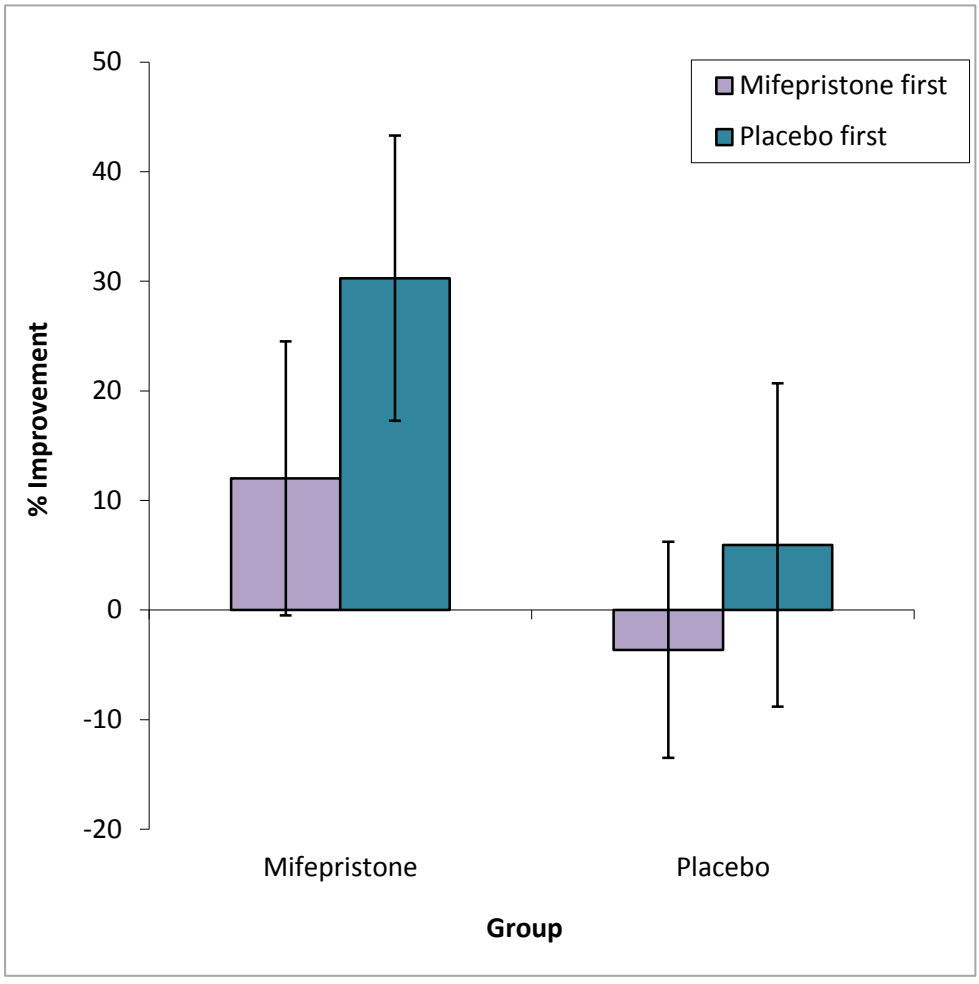
A significant ANCOVA main effect of treatment was found in the SWM between search error (BSE) rate (square root transformed) of the spatial working memory task. Subsequent analysis of this significant main effect revealed that, following mifepristone treatment, the error rate was significantly reduced from baseline (t=3.04, df=18, p=0.007). However no significant change occurred following placebo (t=1.24, df=18, p=0.232). Direct comparison of the treatments revealed a significant advantage of mifepristone over placebo in the percentage improvement (calculated for each individual subject) in error rate from baseline (mean difference=19.8%, 95%CI=4.3 to 35.2; t=2.69, df=18, p=0.015) (see Figure 3-2; upper). Order of treatment administration did not appear to be a confounding factor. The improvement following mifepristone was not significantly different in the group who received mifepristone first compared to the group who received it second. Again there was no difference in the response to placebo between these groups (p>0.2 for all). There was also no ANCOVA main effect of order or treatment by order interaction (see Figure 3-2; lower). There were no significant main effects of treatment on any outcome measure from the Rey-AVLT (total correct, long-term recall or recognition).

Table 3-1. Neuropsychological test results at baseline and following mifepristone and placebo

	Baseline		Mifepristone		Placebo		ANCOVA treatment main effect	
	Mean	s.d.	Mean	s.d.	Mean	s.d.	F(*df*=1,16)	*P*=
PRIMARY MEASURES								
Spatial working memory								
Between search errors	43.32	31.88	34.16	27.06	37.89	22.78	7.12	**0.017**
Within search errors	4.74	13.05	2.00	5.33	3.00	5.55	2.42	0.139
Rey-AVLT Total correct	42.32	10.97	40.68	10.14	43.68	12.85	1.36	0.216
Delayed recall	7.05	3.37	7.05	3.70	6.21	3.26	0.04	0.845
Delayed recognition	11.16	3.40	10.32	3.33	11.37	2.50	0.03	0.859
SECONDARY MEASURES								
Memory span								
Forward digit span	6.42	1.12	6.37	1.26	6.37	1.21	0.01	0.943
Forward spatial span	4.95	1.22	5.11	1.10	5.32	1.25	<0.01	0.963
Executive function								
Verbal fluency (correct)	32.74	11.93	36.79	11.46	35.89	10.3	4.53	**0.049**
Backward digit span	4.47	1.58	4.53	1.39	4.11	1.29	<0.01	0.997
Vigil; Total correct (%)	85.11	24.16	92.42	8.93	90.89	11.31	1.52	0.235
Learning and Memory								
Spatial recognition (errors)	7.42	3.50	6.74	3.21	7.63	2.77	4.99	**0.040**
Pattern recognition (errors)	3.79	3.55	3.89	3.75	3.58	3.15	0.78	0.391
Attention/psychomotor								
Digit Symbol (correct 90s)	43.11	7.18	44.26	10.38	45.89	9.53	2.17	0.161
Vigil; Latency (ms)	402.13	66.03	406.97	65.61	422.22	68.73	0.89	0.360

Figure 3-2. Improvement in SWM between search error rate overall (top) and by order (lower)





3.3.1.2 Secondary outcome measures

ANCOVA main effects of treatment were found in both spatial recognition memory and verbal fluency (see Table 3-2).

For verbal fluency, the number of words correctly produced was significantly greater than at baseline following mifepristone treatment (t=3.34, df=18, p=0.004) with no significant difference following placebo (t=1.57, df=18, p=0.133). Direct comparison of each treatment, expressed as a percentage improvement from baseline, did not significantly differ (mean difference=1.60%, 95%CI= -9.89 to 13.10; t=0.29, df=18, p=0.773). For the spatial recognition task, direct comparison of mifepristone versus placebo, expressed as a percentage change in error rate from baseline, revealed a trend towards a lower error rate following mifepristone (mean difference=27.2%, 95%CI= -1.81 to 56.17; t=1.97, df=18, p=0.064).

3.3.2 Symptoms

At +14 days, following treatment with mifepristone, depression rating scores from the $HDRS_{17}$ and MADRS had significantly improved from baseline levels (see Table 3-2). No significant change was observed at any time point following placebo. Direct comparison of the advantage of mifepristone over placebo at this time point (+14 days), however, failed to reach statistical significance for either $HDRS_{17}$ scores (mean difference=2.32, 95%CI= -2.08 to 6.71; t=1.107, df=18, p=0.283) or MADRS scores (mean difference=2.26, 95%CI= -3.36 to 7.89; t=0.845, df=18, p=0.409).

Table 3-2. Symptom ratings in the group as raw scores and change from baseline with 95%CI

	Raw data				Mean (95%CI) improvement from baseline			
	Mifepristone		Placebo		Mifepristone		Placebo	
	Mean	s.d.	Mean	s.d.	Mean	95% CI	Mean	95% CI
$HDRS_{17}$								
Baseline [a]	18.05	9.92	-	-	-	-	-	-
Day +7	15.42	7.42	14.89	8.31	2.63	-1.50 to 6.76	3.16	-0.37 to 6.69
Day +14	12.95	6.74	15.26	9.33	**5.11** [b]	0.91 to 9.30	2.79	-2.11 to 6.77
Day +21	16.68	8.27	14.58	7.60	1.37	-3.23 to 5.97	3.47	-0.86 to 7.81
MADRS								
Baseline [a]	23.00	10.24	-	-	-	-	-	-
Day +7	20.58	9.96	20.16	11.08	2.42	-2.62 to 7.46	2.84	-0.49 to 6.18
Day +14	16.95	10.56	19.21	10.65	**6.05** [c]	0.75 to 11.36	3.79	-0.42 to 8.00
Day +21	20.37	10.89	19.79	10.17	2.63	-2.93 to 8.20	3.21	-2.33 to 8.75
BPRS								
Baseline [a]	29.79	7.14	-	-	-	-	-	-
Day +7	26.95	7.38	27.63	6.40	2.84	-1.03 to 6.72	2.16	-1.00 to 5.32
Day +14	25.74	7.99	26.84	6.03	**4.05** [d]	0.12 to 7.99	2.95	-1.17 to 7.06
Day +21	26.79	6.44	26.68	6.81	3.00	-0.73 to 6.73	3.11	-1.46 to 7.67
YMRS								
Baseline [a]	3.58	3.78	-	-	-	-	-	-
Day +7	3.37	4.62	4.16	4.05	0.21	-1.97 to 2.39	-0.58	-2.26 to 1.10
Day +14	2.89	3.63	4.11	4.08	0.68	-1.37 to 2.74	-0.53	-2.79 to 1.74
Day +21	3.00	2.47	3.21	2.97	0.58	-0.78 to 1.94	0.37	-1.63 to 2.37

[a] *Baseline for both treatments (mifepristone and placebo)*
[b] *Significant change from baseline (t=2.56, df=18, p=0.020)*
[c] *Significant change from baseline (t=2.40, df=18, p=0.028)*
[d] *Significant change from baseline (t=2.16, df=18, p=0.044)*

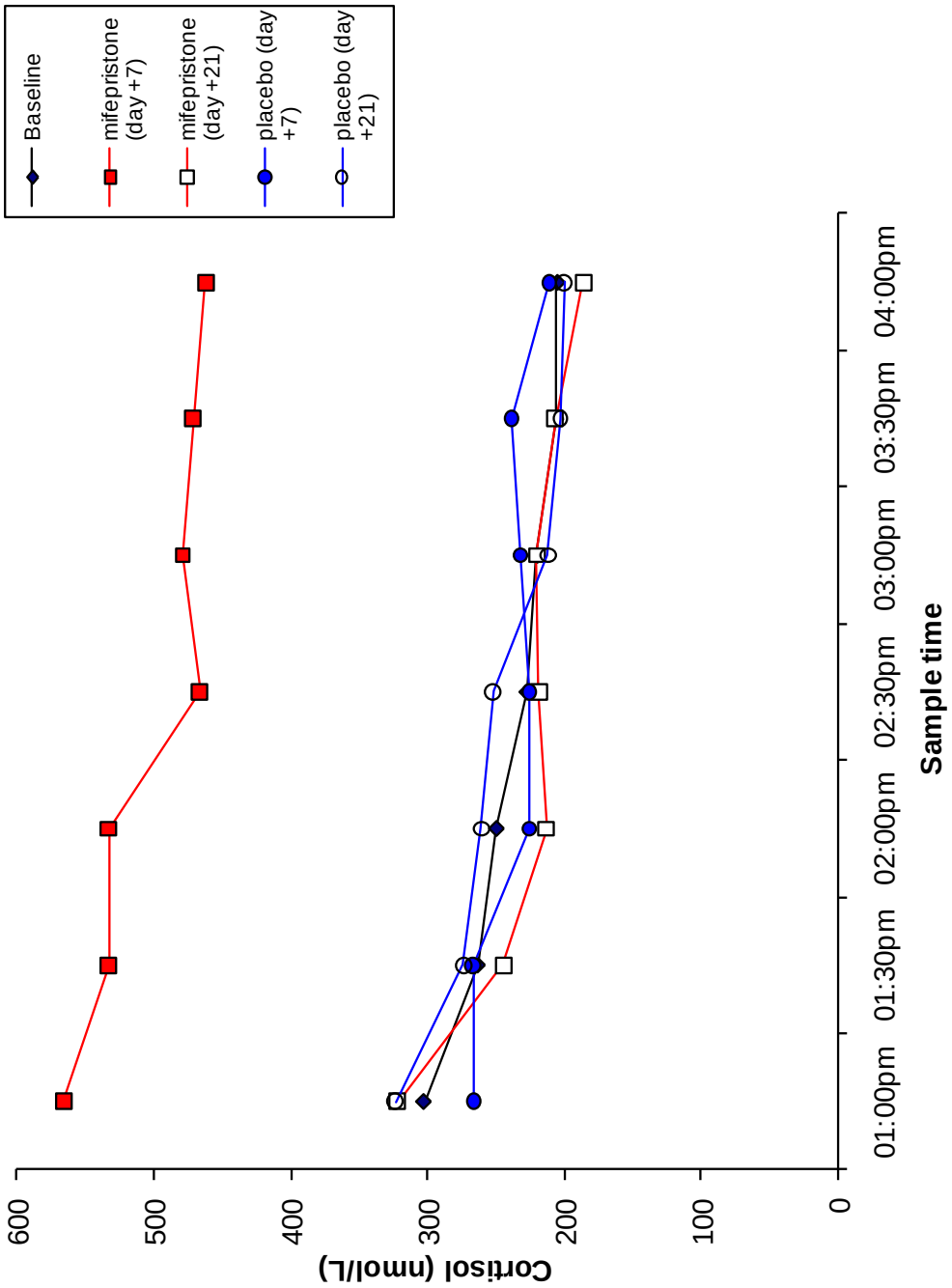
An independent samples t-test was used to confirm that the order of treatment administration was not a confounding factor. There was no significant difference in response to the active treatment, between the group receiving mifepristone first or the group receiving it second in either $HDRS_{17}$ scores (t=0.054, df=17, p=0.958) or MADRS scores (t=0.554, df=17, p=0.587).

Of the secondary scales, BPRS scores were also found to be significantly lower at +14 days following mifepristone treatment, with no significant change following placebo (see Table 3-2). Again however, comparison of the advantage of mifepristone over placebo at this time point failed to reach statistical significance (mean difference=1.11, 95%CI= -3.00 to 5.22; t=0.564, df=18, p=0.579). YMRS scores did not significantly differ from baseline at any time point. A *post hoc* analysis was performed on all symptom effects, after the exclusion of the three patients who did not fulfil SCID criteria for a current depressive episode. The improvement from baseline at +7 days remained significant for all measures (p<0.05).

3.3.3 Neuroendocrine measures

A highly significant ANOVA main effect was observed (F=20.6, df=4,68, p<0.0001), with cortisol levels being significantly higher following mifepristone treatment (day +7) compared to all other visits (see figure 3). A significant diurnal rhythm was evident in the effect of time (F=21.6, df=6,102, p<0.0001), although there was no interaction between visit and time (F=1.18, df=24,408, p=0.29). No other significant effects were observed.

Figure 3-3. Cortisol levels 1pm to 4pm at baseline and following mifepristone or placebo at day +7 and +21

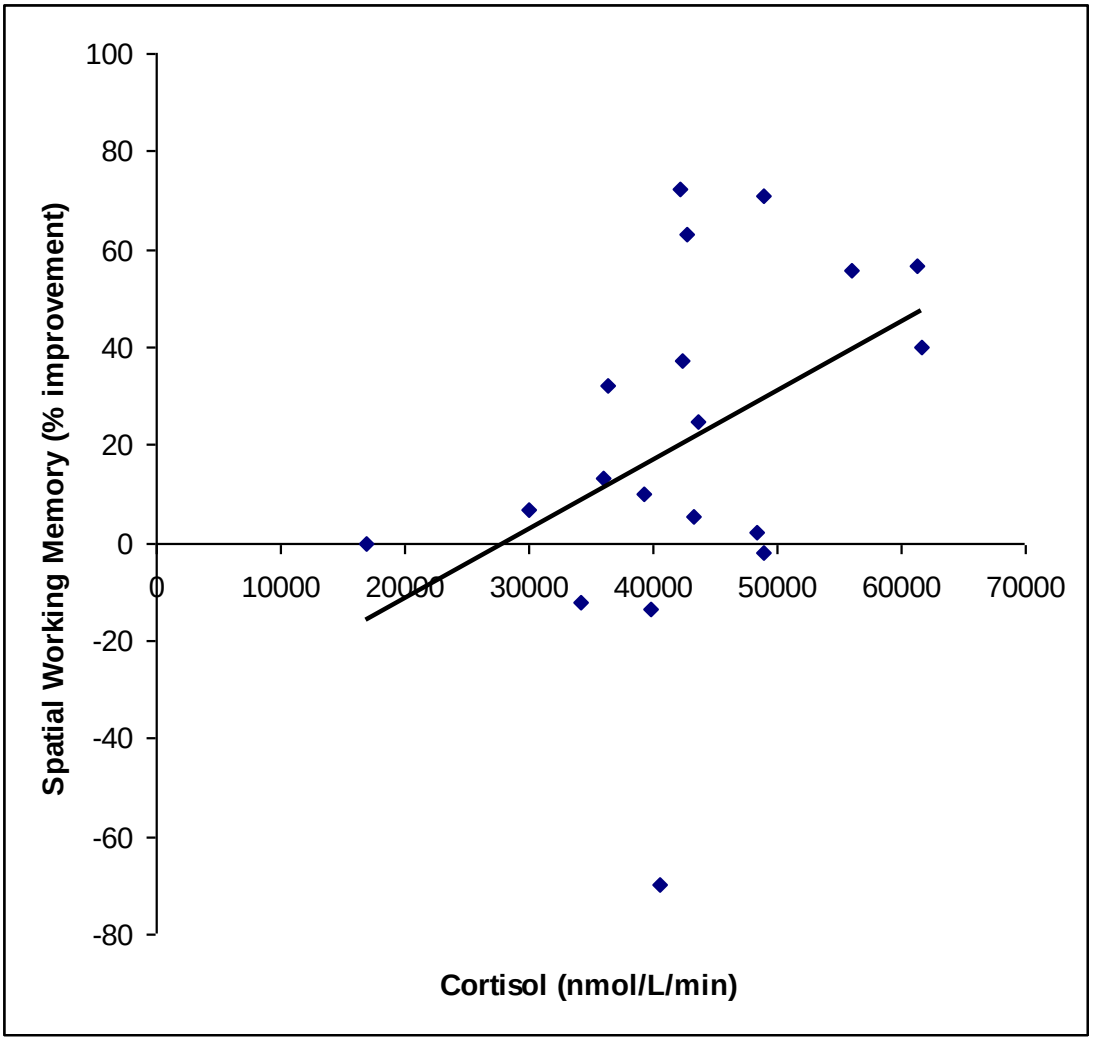


Subjecting the data to the same analysis procedure as the neuropsychological tests (the method recommended byVickers & Altman, 2001), the AUC cortisol data from day +21 (the day of neuropsychological testing) was entered into a repeated measures ANCOVA with baseline and age as covariates. A significant main effect of treatment was observed, with cortisol levels being significantly lower after mifepristone treatment (mean AUC=40,673, s.d.=11,266) compared to placebo (mean AUC=44,046, s.d.=9,884; F=6.449, df=1,16, p=0.022).

3.3.4 Exploratory analysis: relationship between cortisol and neuropsychological functions

An exploratory *post hoc* analysis (see Figure 3-4 below) revealed that the area-under-the-curve (AUC) cortisol output at baseline correlated positively with the percentage improvement in spatial working memory error rate following mifepristone administration (r_s=0.460, p=0.048). No relationship was found between cortisol AUC and the error rate following placebo (r_s=0.286, p=0.235).

Figure 3-4. Correlation between baseline cortisol (AUC: 1pm to 4pm) and spatial working memory improvement following mifepristone



Further exploratory analysis was conducted to see if there was any relationship between the percentage change in cortisol AUC and change in neuropsychological test score, in other words, does the extent of the suppression of cortisol by mifepristone correlate with the degree of improvement in neuropsychological function? From the result of the above graph (Figure 3-4), the one outlier was excluded prior to the analysis. Where appropriate, the sign of the direction of the correlation coefficient was reversed so that all values are consistent, irrespective of whether the outcome measure uses errors or correct responses (i.e. negative correlations are expected in the mifepristone arm, so that a greater reduction in cortisol levels

is associated with a greater improvement in neuropsychological function). From the table below, the correlation between the primary outcome measure of Spatial Working Memory was at trend level only (r_s= -0.355, p=0.087), however consistent significant results were observed for Spatial Recognition Memory, delayed verbal recognition and psychomotor speed (DSST). All were consistent with the hypothesis that a great suppression of cortisol would be associated with a greater improvement in neuropsychological function (It should however be noted that one significant correlation – Vigil commission errors – went contrary to this prediction).

Table 3-3. Spearman's correlations between neuropsychological improvement and cortisol suppression in response to mifepristone (both examined as the percentage change)

	Mifepristone	Placebo
SWM between errors	-0.335	-0.172
Spatial Recognition	**-0.458***	0.104
Verbal fluency (FAS)	0.120	0.251
DSST	**-0.459***	-0.135
Pattern Recognition	0.063	0.262
SSP	0.270	0.050
Digit span forward	0.137	-0.152
Digit span reverse	-0.238	-0.084
Vigil omissions	0.344	-0.229
Vigil commissions	**0.572***	0.075
Stroop CW	-0.279	0.118
Rey total A1 to A5	-0.158	-0.295
Rey A7	-0.169	-0.292
Rey A recognition	**-0.459***	-0.093

** p<0.05; Directional hypothesis (one-tailed)*

3.4 Discussion of Chapter 3

These data suggest that administration of a drug with GR-antagonist properties – mifepristone – selectively improved neuropsychological functioning in bipolar depression. Spatial working memory function (between search errors) was significantly improved compared to placebo (measured 2 weeks after cessation of treatment). Changes in secondary measures of spatial recognition memory and verbal fluency were observed, but these improvements were not significant over those seen with placebo. Ratings of depression ($HDRS_{17}$ and MADRS) and total BPRS scores were similarly reduced compared to baseline after treatment with mifepristone, but not after treatment with placebo (this occurred 7 days after cessation of treatment). There was no significant difference from baseline in mood ratings at the time of neuropsychological testing (14 days after cessation of treatment). Analysis of the neuroendocrine data revealed that cortisol levels in the group had significantly reduced from baseline at this point. Exploratory correlational analyses revealed that the extent of the reduction in cortisol levels correlated with the percentage improvement in some of the neuropsychological measures (see Table 3-3). While this was only at a trend-level with the primary outcome measure – SWM – it was found that cortisol levels at baseline correlated with the extent of the improvement in this measure (Figure 3-4).

As noted in the introduction, animal work has shown that administration of a GR-antagonist can, under certain treatment protocols, improve spatial memory (Oitzl & de Kloet, 1992; Oitzl et al., 1998a; Oitzl et al., 1998b). Chronic administration was shown to improve spatial memory as did direct administration to the hippocampus. The relatively high dose of mifepristone used in the present study, along with the long half-life of the drug, may have resulted in chronic blockade by the end of the treatment period. However, because of the crossover design of the study and the delayed neuropsychological testing period (which was 2 weeks after cessation of treatment) it is not possible to ascertain if the improved spatial

memory processes in the present study are directly attributable to this effect. An alternative possibility is that the improvement is related to the overall suppression or 'normalisation' of the HPA axis. One previous study has shown that a single dose of mifepristone (200mg) can lead to a reduction in glucocorticoid bioactivity for up to 2 weeks after the treatment (Heikinheimo et al., 2003b). The correlations (whilst being moderate and only obtained through exploratory analyses) between the extent of cortisol suppression after mifepristone and the degree of improvement in neuropsychological functioning may suggest that this direct link also contributes to the effect. It should however be noted that the actions of this class of drug are highly complex and may be attributable to other factors outside of those that can be ascertained from the data here. For example, although it is typically referred to as an antagonist, evidence indicates that mifepristone is a partial agonist at the GR (Bourgeois et al., 1984; Laue et al., 1988). A recent paper has reported that mifepristone exerted partial agonistic effects, while blocking the effects of glucocorticoids, on mitochondrial GR translocation and mitochondrial membrane potential. These results are discussed the context of the biphasic (inverted "U"-shaped) effects of glucocorticoids on neural functions, including memory (Du et al., 2009). Interestingly, in a recent animal study, of the series of GR antagonist examined, mifepristone was the only drug to increase both mineralocorticoid receptor (MR) and GR binding in the frontal cortex (Bachmann et al., 2003) which may lead to specificity in the effects on neuropsychological functions seen here (see next section 3.4.1).

It is of course possible that the apparent incongruence in the results in other domains (such as verbal declarative memory) compared to previous studies may be attributable to methodological differences. For example, when examining the effects of corticosteroids in the healthy human literature, studies typically use paragraph recall (with its incorporated sentence structure) or other word lists with a larger number of items. The present study found no effects of mifepristone on verbal declarative memory which used the 15-item Rey-AVLT,

while subtle but positive effects were seen on word recall in the Pomara et al study (although it should be noted that the 'trend' described by the authors would not meet most definitions used by others) (Pomara et al., 2002). It is also important to consider that the tasks on which improvements were observed were associated with the largest effect sizes within their domain in Chapter 2. Verbal fluency and SWM between search errors were the largest effect sizes within the Executive domain, and Spatial Recognition the largest within the Visuo-spatial memory domain (and also met criteria to be classified as a differential deficit) and therefore it could be argued that effects were relative to the size of the deficit to begin with[6]. However, other processes such as psychomotor speed (DSST) were not affected and yet exhibited large effect sizes, so unless the effect is process specific and effect size sensitive, then this cannot explain the result wholly.

3.4.1 Implications of the preceding findings for the subsequent direction of the thesis

So as a starting point for the subsequent studies in this thesis, let us take the position that there may relative specificity for the effects of HPA axis manipulation on neuropsychological functioning. The tasks affected to the greatest extent following direct GR manipulation was CANTAB Spatial Working Memory (SWM), along with trends (i.e. significant change from baseline, but not over placebo) in Spatial Recognition and Verbal Fluency. Therefore it is worth considering in more detail the processes and brain structures underlying these tasks.

[6] Although note that the notion of 'regression towards the mean' is not relevant here as the crossover design would make it equally as likely that the lower scores occur in the placebo arm of the study as in the active arm.

Focussing first on the CANTAB SWM test, a number of studies have developed variants of this paradigm such as the Executive Golf Test (Feigenbaum et al., 1996; Morris et al., 1996) and the boxes task (van Asselen et al., 2005; van Asselen et al., 2006). All involve self-ordered searching across fixed locations, with subjects having to remember accumulating locations of positions where targets have been found and avoid repetitious searching. Consequently, outcome measures include within search errors (returning to a location already searched within a given search sequence) and between search errors (returning to a location where a target was already found previously). It has been suggested that the latter reflects the maintenance of information held in memory over a longer time period, while the former can be considered to be more immediate and a measurement of the visuo-spatial sketchpad of working memory (Feigenbaum et al., 1996; van Asselen et al., 2005; van Asselen et al., 2006). Generally such tasks are considered to broadly tap executive and working memory processes, but may extend beyond this. Early work examining the underlying the underlying neural circuitry of these processes indicated that patients with unilateral or bilateral frontal lesions made significantly more of both types of error than matched controls and these increased with set-size (Owen et al., 1990). In a subsequent study these findings were extended to groups of patients with frontal lobe lesions, temporal lobe or following unilateral amygdalo-hippocampectomy (AH) where it was reported that significant impairment in between search errors[7] was not observed in the temporal group, but was in the other two, even at the lowest levels of task difficulty in the case of the frontal group (Owen et al., 1995). However it should be noted that the temporal group did make more errors at trend level and in a subsequent study using the 'Executive Golf' variant of the task, significantly increased between search errors have been reported with both unstructured and structured (by eliminating the self-ordered search element of the task) versions of the test in the absence of differences in within-search errors (Feigenbaum et al., 1996). Therefore, as has been suggested (see van

[7] Within-search errors were not reported in this study.

Asselen et al. 2005, 2006), it does appear that the longer-term maintenance of spatial information, reflected in the between search error rate, is reliant upon structures like the temporal lobes and hippocampus, while within search processes are unaffected unless there is frontal lobe impairment. This general position was supported in a recent study in stroke patients (van Asselen et al., 2006) which looked in more detail at the brain areas underpinning spatial working memory performance, both in terms of lesion location and general hemispheric specialisation. Results indicated there were no differences in immediate spatial memory with the Corsi-blocks task, both in terms of specific location or hemispheric specialisation, although damage to the right DLPFC and the right PPC was correlated with performance. On the spatial working memory test ('boxes task'), within search errors were significantly higher overall in right-hemisphere (RH) patients than controls while no differences were found between left hemisphere (LH) patients and the control group. Analysis of specific brain areas (dorsolateral prefrontal cortex, DLPFC; posterior parietal cortex, PPC; hippocampal formation, HF) revealed no significant difference between RH, LH and controls within any of these areas[8]. With respect to between search errors, again RH patients made significantly more than controls while there was no significant difference between LH and controls. Location specific analyses revealed that right DLPFC performed worse than controls, with no difference between left DLPFC patients and controls. A similar pattern emerged with regard to the PPC, where RH performed worse than controls. For the HF, both RH and LH patients performed worse than controls. Overall the results were indicative of a general involvement of the RH in spatial working memory processes; the authors suggested specifically that the DLPFC and the PPC may be essential for keeping spatial information in memory over short periods, while the HF is involved in the transfer of information from working memory into long-term memory or vice versa i.e. accessing long-term memory for transfer to working memory, to facilitate performance (van Asselen et al., 2006).

[8] However, it should be noted that the overall rates of within search errors tend to be minimal compared to between search errors and therefore observing significant changes is often difficult.

The original question of why these tests specifically were improved by GR manipulation could perhaps be related to the actions of the drug or cortisol within these areas, especially if we tentatively consider as supportive the other tests that showed significant improvement from baseline. The Spatial Recognition Memory task was also examined in the study by Owen et al. (1995) and performance was found to be significantly worse in the 'frontal lobe group' only, with no significant effect in the AH or temporal groups (although it should be noted that the actual level of significance in the latter was p=0.075 and therefore a trend at least is suggested). With regard to verbal fluency, many studies have demonstrated the involvement of the frontal cortex in phonological fluency tasks (Davidson et al., 2008), particularly the LH (associated with semantic retrieval) with some temporal lobe activation (Cabeza & Nyberg, 2000; Weiss et al., 2003). Again these tests are very sensitive to frontal damage but are also significantly impaired in many patients with temporal or hippocampal damage (Gleissner & Elger, 2001; Alessio et al., 2006). Therefore, is it the case that the tasks which are altered through HPA axis manipulation reflecting the direct, underlying actions of the drug or are they simply those that engage the broadest range of processes/systems?

As discussed in section 3.1.6.2.1 above, administration of mifepristone in rodents can significantly improve spatial memory when administered chronically and when administered directly to the hippocampus. Also, we have noted the drug has been shown to increase corticosteroid receptor numbers in the frontal lobes. Therefore there is some face validity at least to this assertion although stronger evidence is needed, particularly as our understanding of the complexity of human memory functions (and underlying theoretical models) has developed. To some, especially within the human experimental psychology literature, we have moved away from equating and classifying tasks with specific brain regions. For example, this is illustrated well in the move away from discussing 'frontal' tasks and instead focussing on underlying executive processes (Baddeley & Wilson, 1988). And it is this 'process' approach

that is needed to continue the exploration of these results. To better understand the specificity of effects, especially the spatial aspect, we need first to examine the factor structure underlying the tests used: what are the shared factors or common processes involved? This is important for the SWM because whilst it has been used in some studies in bipolar depression (see section 1.3.2) and in those studies described above linking brain structure to the test measures, there is a lack of understanding of the neuropsychological processes underpinning performance on SWM. From this, we can examine a number of questions that have emerged from the data so far: what are the factors underpinning the SWM task, specifically the between search errors? Is the factor structure in the patients different to that of healthy controls i.e. are some processes scaffolded by others?[9] Secondly, to subsequently explore the spatial effect we first need to consider the processes underpinning the general concept of 'spatial memory' and attempt to utilise tests that allow fractionation of these. Thirdly, we should extend the measures of HPA axis function to include measures more closely related to the function of the GR receptor.

This discussion above frames the subsequent three chapters of the thesis. In Chapter 4, Principal Component Analysis (PCA) will be employed to examine the factor structure of neuropsychological processes in a larger group of bipolar depressed patients and healthy controls. In Chapter 5, the results of a novel spatial (object-location) memory task will be reported which permits fractionation of distinct processes. Finally, in Chapter 6, the relationship between distinct spatial measures, general neuropsychological factors and newer more sensitive HPA axis measures are explored (specifically, the spatial measures of interest are those of the SWM, because of the findings in Chapter 3 where effects emerged in the between but not within search errors; and also the novel task described in Chapter 5),

[9] This has been noted in many different contexts. For example, using PET imaging, very different patterns of activation have been observed in patients with Parkinson's Disease who were able to maintain performance on an executive planning task through a shift to a declarative memory processing /'hippocampal' activation (Dagher et al., 2001). The issue of executive scaffolding of visuospatial WM has also been demonstrated in euthymic bipolar patients (Thompson et al., 2006).

Chapter IV

The component structure of neuropsychological processes in bipolar depression and healthy controls

4. The component structure of neuropsychological processes in bipolar depression and healthy controls

4.1 Introduction to Chapter 4

As outlined in the previous discussion the purpose of the present chapter is to analyse in detail the component structure [10] of neuropsychological processes in bipolar disorder. The tests examined will include those that have been used in chapters 2 and 3 of this thesis alongside some additional measures to tap specific processes not covered earlier or to address methodological issues that had arisen. These are briefly discussed below.

4.1.1 Additional neuropsychological tests for PCA

One issue that can occur with the use of some CANTAB tests in healthy or high-functioning participants is that of ceiling effects (see section 2.3.2.2). In Chapter 2, 8 participants (20%) achieved the maximum score of 24 on the Pattern Recognition (PRec) test and therefore the observed effect may have been an underestimate (as this occurred in 5 controls and 3 patients). There are two likely explanations for this: firstly, the patterns used in the task in its present form are highly nameable and loosely resemble many concrete objects; and secondly, the test is administered in two blocks of twelve items. A modified PRec (PRec-m) test was therefore developed using the more abstract, black-and-white patterns (Vanderplas & Garvin, 1959) and administered as a single block of 24. All other parameters remained the same as the original version. Although no participant achieved a maximum score on the Spatial Recognition (SRec) test, a number of people came close and therefore in addition to the standard test of four trials with 5 spatial locations to remember, two additional modified trials

[10] This will be carried out using Principal Component Analysis (PCA). The specific methodology used here will be outlined subsequently.

with 7 locations (SRec-m7) and two with 9 locations (Srec-m9) were introduced for use with this larger cohort.

In this study, the reverse version of the CANTAB Spatial Span test was also included to parallel the testing procedure used in the digit span test. The 'forward' versions of the digit and spatial span tests broadly tap the phonological loop and visuo-spatial sketchpad respectively, while the reverse versions place greater demands on the central executive component of working memory. A modified verbal fluency test (Crawford et al., 1995; Shores et al., 2006) was also included – the 'Exclude-Letter' Fluency Test (ELFT) – which again places greater demands on executive processes. Unlike the standard 'FAS' test in which subjects are required to produce as many words as possible beginning with a given letter, in the ELFT participants must produce as many words as possible that *do not* contain the given letters (vowels are typically used). To retain the characteristics of the FAS version the time for each letter was reduced to 60s (rather than the 90s as described in the original study). Lastly, an additional measure of speed of cognitive processing (SCOLP; Speed and Capacity of Language Processing) was included. This test more specifically measures slowing of language processing (Baddeley et al., 1992).

Additional tests were included to tap specific aspects of (visuo-spatial) working memory and its executive control. The abstract design version (Petrides & Milner, 1982) of the Self-Ordered Pointing Test (SOPT) is a test of executive function/ visual working memory, requiring the ability to generate and monitor a sequence of responses. As the patterns used are abstract and change spatial position after each trial, participants must rely on visual-strategic processing in the task. The version used here consists of 3 trials at levels 4, 6, 8 and 10. A second visual memory test was included – the Visual Patterns Test (VPT) – which was designed to measure 'short-term visual memory largely devoid of its spatio-sequential component' [11]

[11] This is a direct contrast to the Corsi or Spatial Span test which is a purer measure of spatio-temporal processing.

(Della Sala et al., 1997; Della Sala et al., 1999). The version used in this study was computerised, with presentation of the stimuli on a PC but responses were recorded on paper as per the original test. In order to more closely parallel the spatial span test, the VPT allows 3 attempts at each level before terminating the test.

The addition of these tests will permit a more detailed interpretation of the factors that emerge from both the PCA and when used as composite scores in subsequent regression analysis.

4.1.2 Aims of the PCA

There are two broad aims of the present chapter and PCA. Firstly, using all the tests available (as outlined), a PCA will be carried out on a larger sample of patients with bipolar disorder and healthy control participants. The aim here is not to repeat the study reported in Chapter 2 (i.e. examining generic performance differences between patients and controls) but instead to explore the common factors onto which the different tests and processes load. Secondly, as already discussed, one of the aims of the subsequent chapters is to obtain a better understanding of the specificity of effects on spatial memory processes, therefore the PCA will also be run with a test of interest, such as the CANTAB Spatial Working Memory, removed in order to establish unique factor scores/unweighted composites that can be used in the analyses in the final empirical Chapter 6. In that final chapter, the composite scores derived from the results of this PCA and additional HPA axis measures will be used in regression analyses to explore the relationship with the primary spatial measures (namely, Spatial Working Memory and the Object-Location Memory paradigm in Chapter 5).

4.2 Methods

For this and all subsequent chapters, a new cohort of patients and controls was recruited as part of an extended research programme into the effects of GR antagonists in bipolar depression funded by the Stanley Medical Research Institute (SMRI) and the Medical Research Council (MRC).

4.2.1 Subjects

Patients (n=53) 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. 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 standardized 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 (n=47) were recruited by advertisement and from hospital/university staff. All controls had also been through a screening procedure prior to testing to exclude anyone with a personal or family history (first-degree) of psychiatric illness, significant medical or neurological illness likely to affect neuropsychological functioning, or history of drug/alcohol abuse.

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.

4.2.2 Neuropsychological tests

As discussed in the introduction, in this part of the study an extended neuropsychological test battery was employed, with particular focus on additional tests of visual and spatial memory. These have been utilised in previous studies and are briefly listed below.

CANTAB Spatial Working Memory (SWM): a self-ordered search task which requires subjects to search for hidden tokens within a spatial array. The number of between-search errors are recorded i.e. occasions when a subject returns to a square under which a token was already found, as well as a strategy measure, were a lower strategy score reflects a more systematic search strategy.

CANTAB Spatial Recognition (SRec): a memory task in which subjects view 5 'squares' presented in serial order and then are subsequently required to identify, from a choice of 2 squares, the one that occupies one of the 5 locations shown previously. Subjects complete 4 sets. The percentage of correct responses are recorded.

CANTAB Spatial Recognition-modified (SRec-m): a modified version of the task was also administered which is identical to the standard version except two sets of 7 squares, then 2 sets of 9 squares are used.

CANTAB Spatial Span and Reverse Spatial Span (SSp/ rSSp): a test analogous to the Corsi Block task which is administered first in the standard format and then reverse, where subjects tap the sequence in the opposite order from presentation. The maximum span reached is recorded.

Visual Patterns Test (VPT): a test of short-term visual memory in which subjects are required to remember and reproduce increasingly complex 'checkerboard' patterns (Della Sala et al., 1999). It is scored in the same way as the SSp task with the maximum set-size achieved being recorded.

CANTAB Pattern Recognition (PRec): a test of visual recognition memory in which subjects view a series of 12 coloured patterns and must then select the patterns they have seen in a 2-choice, forced-discrimination paradigm. Subjects complete 2 sets and the percentage correct is recorded.

Pattern Recognition-modified (PRec-m): due to the risk of ceiling effects in healthy controls, a modified pattern recognition task was constructed which was similar to the CANTAB version except the patterns were more abstract, black and white shapes and were more closely matched to their distracter during the recognition phase. These were taken from (Vanderplas & Garvin, 1959) and displayed using the Superlab program. One set of 24 patterns was administered.

Self-Ordered Pointing Test (SOPT): a test of visual memory and strategic processing, using set sizes 4, 6, 8 and 10.

The other tests included are:

Rey-Auditory Verbal Learning Test (Rey-AVLT): a verbal learning and memory task which was administered according to standardised instructions (Rey, 1964; Lezak et al., 2004).

Forward and Backward Digit Span (fDSp/ bDSp): a test of immediate verbal recall and working memory which was again administered according to standardised instructions (Lezak et al., 2004).

Verbal fluency (FAS) and *Exclude-Letter Fluency test (ELFT):* tests of executive function in which participants are required to produce as many words as possible beginning with, or not containing a given letter.

Digit Symbol Substitution Test (DSST): as a test of psychomotor speed and attention.

Speed and Capacity of Language Processing (SCOLP): to test the speed and efficiency of cognitive processing.

4.2.3 Statistical analysis procedure

Due to the exploratory nature of this type of analysis procedure, the general methodology and data screening considerations are outlined in this section. The approach adopted follows closely the recommendations by Field and Stevens (Stevens, 2002; Field, 2009).

4.2.3.1 Preliminary data cleaning

A Principal Component Analysis (PCA) was performed on the neuropsychological tests described above. Of the tests included, on those that have multiple possible outcome measures associated with them only the most commonly used / most representative of specific cognitive processes were extracted: this resulted in 26 potential variables. These 26 variables were then assessed along a number of criteria for inclusion in the analysis. First, as outlined in the introduction, a number of experimental measures were introduced to assess and overcome the issue of ceiling effects in the Pattern Recognition (PRec) and Spatial Recognition (SRec) tests. In the standard version of PRec, n=18 (22.2%) of participants achieved the maximum possible score on the task, therefore this variable was excluded in favour of the modified PRec test where only 2 participants (1.6%) scored the maximum. With the SRec, the standard version of the test was retained as only n=2 (1.6%) of participants achieved the maximum score and SRec 7 and 9 were excluded.

Test variables were also assessed for missing values: all those retained had a maximum of n=5 (5%) missing data points, which were replaced with the group mean (patient or control).

However the 3 variables from the Vigil test were excluded as, due to a technical error in the software, only n=75 (75%) valid data was available.

Finally, as recommended by Field, the correlation matrix was examined for any extreme values i.e. variables correlating very highly or very weakly with others. For the Rey-AVLT, delayed recall of List A (trial A7) unsurprisingly correlated very highly with A7 percentage retained ($r>0.9$) therefore the latter variable was excluded in favour of the former which showed moderate, significant correlations with a greater number of other variables.

By excluding these outcome measures/variables there are 19 potential for inclusion in the PCA. This results in around 5.3 cases per variable which falls just within the recommended 5 to 10 per variable with samples n<300 (Kass & Tinsley, 1979). This is only a general guideline and more formal testing of the sample and the data will be performed through the iterative process of extracting stable factor solutions. The Kaiser-Mayer-Olkin measure of sampling adequacy (KMO) will be reported along with Haitovsky's test for multicolinearity:

$$\text{Haitovsky's } \chi^2_H = \left[1 + \frac{(2p+5)}{6} - N\right] ln(1 - |\mathrm{R}|)$$

Where *p* is the number of variables in the correlation matrix, N is the overall sample size and $|\mathrm{R}|$ is the determinant of the correlation matrix. The resulting Chi-squared statistic has degrees of freedom:

$$\frac{p(p-1)}{2}$$ (Haitovsky, 1969; Field, 2009).

A significant result on this test indicates there is no severe multicolinearity (although it should be noted, that assumption is not necessary with PCA). As recommended by Stevens, with samples of around n=100 it is also recommended that a significant result on Bartlett's test of

sphericity is obtained in order to *reject* the null hypothesis that the variables in the population correlation matrix are uncorrelated.

4.2.3.2 Factor rotation selection

Finally, an important consideration is the interpretation of the components which are extracted, which is done through the rotation of factors. There are numerous methods available, many of which are included in statistical analyses packages. However, there is considerable debate as to the method to use. The main consideration is whether to select Orthogonal (varimax) or Oblique (direct Oblimin) rotation [12] (Stevens, 2002; Field, 2009). Avoiding detailed discussion of the mathematics behind each of these, the main distinction is that orthogonal rotations produce factors that are uncorrelated while oblique rotations produce factors that will be correlated to a lesser or greater extent. The varimax method (Kaiser, 1960) is design to produce factors that are as independent as possible, with each factor loading as high as possible on a small number of variables and as little as possible with others, hence interpretation of the resulting factors in the component matrix is easier than with other methods (Stevens, 2002). In contrast, the Oblimin method permits factors to correlate, determined to an extent by the delta value entered at the time. The result of this analysis produces two matrices – a pattern matrix which contains the loadings of each variable to the factor with the influence of the remaining variables partialled out; and a structure matrix which presents the simple factor loadings of each variable. Consequently, if the factor structure is orthogonal these two matrices are the same, however differences between the two (in terms of variable loadings) are of interest as these differences indicate those factors/variable loadings that are largely independent compared to those that load across

[12] There are many additional examples of rotation methods within these categories. Those discussed here are those recommended in each case, based on the consistency and predictability of results and ease of interpretation of the resulting factors.

several factors. When dealing with components of human psychological processes it has been argued that it is unlikely that these are ever truly independent (Field, 2009).

Therefore the approach adopted in the present study is to examine and compare the results of both rotation methods (as recommended by several authors Pedhauzur & Schmelkin, 1991; Stevens, 2002). This fits with the two broad aims of the present chapter; first, to understand the factor structure underlying the tests and processes employed, and second, to produce interpretable composite scores for use in subsequent regression analyses. As has been discussed, the varimax method of producing factors that are as unique as possible will create interesting theoretical distinctions between these components with the caveat that they may not be precisely representative of the complex inter-relationships between processes in human neuropsychology. In contrast, the Oblimin method will provide valuable information on the relative independence of each variable or shared component across the resulting factors, and through the iterative process of re-assessing the strength of these factors as those variables that are multiple-loading are removed, a set of independent factors can be produced (at which point the pattern and structure matrices will be the same and hence orthogonal; replicating the varimax model).

One final point to note surrounds the labels given to the extracted components. These should be viewed as broad, general descriptions only, given the complexity of accurately describing the exact combination of tests or measures within any given component.

4.3 Results

Fifty three bipolar patients (33 male, 20 female) participated in the study. Patients were aged between 22 and 63 years (mean=47 years, s.d.=10) and had a NART estimated IQ of 109 (s.d.=2). There were no current psychotic features in the group and no current diagnosis of substance abuse or dependence. The average (median) age of onset[13] in the group was 24 years (mean=27, s.d.=13). The median number of hospitalizations[14] in the group was 1. Twenty six patients (49%) had previously attempted suicide[10] and 11 (22%) had previously been treated with ECT[15] (>12 months ago).

All patients had persistent depressive symptoms, with all fulfilling SCID criteria for current depressive episode. The median length of current depressive episode[12] in the group was 26 weeks (mean=61.5, s.d.=82.7). Depressive symptoms had a mean score of 28 (s.d.=8) on the Montgomery-Åsberg Depression Rating Scale (MADRS)[16] and of 20 (s.d.=5) on the 17-item Hamilton Depression Rating Scale ($HDRS_{17}$).

The healthy control group (n=47) consisted of 28 males and 19 females. Controls were aged between 18 and 64 (mean=45 years, s.d.=14) and had a NART estimated IQ of 112.5 (s.d.=12). This group was matched to the patient group by sex (χ^2=0.76, *df*=1, *p*=0.783), age (*t*=0.954, *df*=98, *p*=0.343) and NART score[17] (*t*=1.586, *df*=93, *p*=0.116).

[13] Data on 4 patients was missing.
[14] Data on 2 patients was missing.
[15] Data on 3 patients was missing.
[16] Data on 1 patient was missing.
[17] n.b. by accidental omission, the NART was not completed for 2 patients and 3 controls.

4.3.1 PCA #01: Total Sample

The first step in the analysis was to further screen the data for variables for inclusion in the PCA. A correlation matrix was produced for all variables (see Table 4-1) and then sorted into order (Table 4-2), allowing exclusion of those variables that did not correlate strongly with others (shading indicates significant $p<0.05$)[18].

4.3.1.1 Initial model

Although there is no fixed criteria for this screening procedure, for the purposes of the initial analysis, Field and others recommend that variables with low-loadings ($r<0.3$) be excluded. To attempt to improve the explained variance within the PCA a more stringent approach to initial data screening was adopted. Any variables with *overall* or *multiple* low-loadings ($r<0.3$) were excluded. Those with the lowest average loading were: SCOLP spot-the-word test (median $r=0.118$) , SWM within errors (median $r=0.120$), and SCOLP speed of comprehension (median $r=0.222$). These variables also had some of the lowest loadings with other individual variables: SWM had no individual correlations above $r=0.28$, for SCOLP spot-the-word test only 3 individual values were $0.30<r<0.46$); on this criteria Forward Digit Span was also excluded (median=0.233 and only 3 individual values $0.30<r<0.40$).

[18] The need to carry out this additional step is indicated as, if an attempt is made to run the PCA with all variables included, the determinant |R| of the correlation matrix is 0 indicating extreme multicolinearity and the rotated factor solution fails to converge (25 iterations; convergence=0.006).

Table 4-1. Overall correlation matrix

	GROUP	SWM berr	SWM werr	SWM strat	PRec-m	SRec	VPT	SSP	SSPr	Rey total	Rey A7	Rey RecA	Dspan -f	Dspan -r	FAS	ELFT	DSST	SoComp	Sp the word
SWM berr	.153																		.023
SWM Werr	.069	.277																.118	
SWM strat	.178	.710	.222															.460	
PRec-m	.284	.329	.036	.332														.567	
SRec	.152	.415	.033	.378	.230													.118	
VPT	.248	.544	.214	.419	.331	.317												.267	
SSP	.241	.540	.149	.428	.254	.289	.479												
SSPr	.270	.497	.198	.331	.259	.387	.500	.563											
Rey total	.406	.426	.189	.280	.313	.382	.232	.376	.373										
Rey A7	.320	.412	.253	.319	.265	.291	.212	.348	.265	.769									
Rey RecA	.225	.413	.032	.340	.336	.423	.127	.214	.138	.551	.650								
Dspan-f	.372	.233	.007	.071	.088	.152	.233	.251	.391	.253	.095	.020							
Dspan-r	.194	.281	.042	.317	.025	.319	.264	.214	.295	.076	.036	.097	.342						
FAS	.315	.213	.180	.290	.255	.246	.219	.263	.239	.263	.191	.171	.274	.409					
ELFT	.446	.276	.092	.354	.167	.307	.282	.228	.351	.369	.302	.265	.212	.354	.640				
DSST	.345	.425	.120	.229	.291	.276	.400	.377	.402	.444	.334	.230	.254	.281	.291	.469			
SoComp	.459	.170	.029	.144	.152	.257	.247	.160	.222	.251	.111	.089	.258	.188	.493	.551	.567		
Sp the word	.153	.023	.082	.123	.207	.158	-.027	-.066	.028	.045	-.060	.073	.191	.258	.410	.370	.118	.460	
SOPT errors	.197	.537	.126	.491	.439	.433	.475	.413	.402	.459	.484	.473	.255	.352	.298	.266	.267	.118	.023

Shading: p<0.05

Table 4-2. Correlation matrix, ranked by magnitude (grey shading indicates median value)

SWM berr	SWM werr	SWM strat	PRecm	SRec	VPT	SSP	SSPr	Rey total	Rey A7	Rey RecA	Dspanf	Dspanr	FAS	ELFT	DSST	SoCo mp	SpotT W	SOPT errors
0.023	0.007	0.071	0.025	0.033	0.027	0.066	0.028	0.045	0.036	0.020	0.007	0.025	0.171	0.092	0.118	0.029	0.023	0.023
0.153	0.029	0.123	0.036	0.152	0.127	0.149	0.138	0.076	0.060	0.032	0.020	0.036	0.180	0.167	0.120	0.089	0.023	0.118
0.170	0.032	0.144	0.088	0.152	0.212	0.160	0.198	0.189	0.095	0.073	0.071	0.042	0.191	0.212	0.229	0.111	0.027	0.126
0.213	0.033	0.178	0.152	0.158	0.214	0.214	0.222	0.232	0.111	0.089	0.088	0.076	0.213	0.228	0.230	0.118	0.028	0.197
0.233	0.036	0.222	0.167	0.230	0.219	0.214	0.239	0.251	0.191	0.097	0.095	0.097	0.219	0.265	0.254	0.144	0.045	0.255
0.276	0.042	0.229	0.207	0.246	0.232	0.228	0.259	0.253	0.212	0.127	0.152	0.188	0.239	0.266	0.267	0.152	0.060	0.266
0.277	0.069	0.280	0.230	0.257	0.233	0.241	0.265	0.263	0.253	0.138	0.191	0.194	0.246	0.276	0.276	0.160	0.066	0.267
0.281	0.082	0.290	0.254	0.276	0.247	0.251	0.270	0.280	0.265	0.171	0.212	0.214	0.255	0.282	0.281	0.170	0.073	0.298
0.329	0.092	0.317	0.255	0.289	0.248	0.254	0.295	0.313	0.265	0.214	0.233	0.258	0.263	0.302	0.291	0.188	0.082	0.352
0.412	0.120	0.319	0.259	0.291	0.264	0.263	0.331	0.369	0.291	0.225	0.233	0.264	0.263	0.307	0.291	0.222	0.118	0.402
0.413	0.126	0.331	0.265	0.307	0.282	0.289	0.351	0.373	0.302	0.230	0.251	0.281	0.274	0.351	0.334	0.247	0.123	0.413
0.415	0.149	0.332	0.284	0.317	0.317	0.348	0.373	0.376	0.319	0.265	0.253	0.281	0.290	0.354	0.345	0.251	0.153	0.433
0.425	0.180	0.340	0.291	0.319	0.331	0.376	0.387	0.382	0.320	0.336	0.254	0.295	0.291	0.354	0.377	0.257	0.158	0.439
0.426	0.189	0.354	0.313	0.378	0.400	0.377	0.391	0.406	0.334	0.340	0.255	0.317	0.298	0.369	0.400	0.258	0.191	0.459
0.497	0.198	0.378	0.329	0.382	0.419	0.413	0.402	0.426	0.348	0.413	0.258	0.319	0.315	0.370	0.402	0.459	0.207	0.473
0.537	0.214	0.419	0.331	0.387	0.475	0.428	0.402	0.444	0.412	0.423	0.274	0.342	0.409	0.446	0.425	0.460	0.258	0.475
0.540	0.222	0.428	0.332	0.415	0.479	0.479	0.497	0.459	0.484	0.473	0.342	0.352	0.410	0.469	0.444	0.493	0.370	0.484
0.544	0.253	0.491	0.336	0.423	0.500	0.540	0.500	0.551	0.650	0.551	0.372	0.354	0.493	0.551	0.469	0.551	0.410	0.491
0.710	0.277	0.710	0.439	0.433	0.544	0.563	0.563	0.769	0.769	0.650	0.391	0.409	0.640	0.640	0.567	0.567	0.460	0.537
0.412	**0.120**	0.319	0.259	0.291	0.264	0.263	0.331	0.369	0.291	0.225	0.233	0.264	0.263	0.307	0.291	**0.222**	**0.118**	0.402

No shading: variables with low-loadings (r<0.3); grey shading: average r value

This resulted in the initial entry of 15 variables into the PCA. The initial model, following factor rotation indicated that the modified Pattern recognition (PRec-m) test did not load onto any component above the pre-defined criteria (see below) and also displayed low communality (0.307). Therefore this variable was eliminated to produce the final PCA reported below using the 14 variables remaining. The factorability of the variables was confirmed using the criteria described above (section 4.2.3). All variables correlated with at least five others at r>0.30 (and less than r=0.77). The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.833 (the cut-off for a 'very good' value is above 0.8) and Bartlett's test of sphericity was significant (χ^2=636.8, df=91, p<0.0001). The diagonals of the anti-image correlation matrix were all much greater than 0.5 (the lowest value was 0.744), justifying the inclusion of each item in the analysis. Although not required for PCA, Haitovsky's test failed to reach significance (χ^2_H =0.09, df=91, p=0.99), however the determinant of the initial correlation matrix was |R|=0.001 (well above the recommended 0.00001) suggesting that multicolinearity is not an issue with these data. Finally, the communalities for the PCA ranged from 0.468 to 0.778, with a mean of 0.690 (see Table 4-3 below).

Table 4-3. PCA #01 communalities (initial model)

	Initial	Extraction
SWM between errors	1.000	0.746
Spatial span	1.000	0.637
Spatial span reversed	1.000	0.653
Visual Patterns test span	1.000	0.644
SWM strategy score	1.000	0.667
SRec Correct	1.000	0.468
SOPT total errors	1.000	0.641
DSST	1.000	0.644
Rey total A1-5	1.000	0.815
Rey A7	1.000	0.827
Rey Recognition A	1.000	0.778
Digit Span (reverse)	1.000	0.637
FAS correct	1.000	0.731
ELFT correct	1.000	0.772

Four factors were extracted, with each independently explaining 40.3%, 11.7%, 9.6% and 7.4% of the variance (cumulatively: 40.3, 52.0, 61.2 and 69.0%).

Table 4-4. PCA #01 unrotated component matrix (initial model) a

	Component			
	1	2	3	4
SWM between errors [b]	-0.777	-0.024	0.360	0.113
Spatial span	0.659	0.132	-0.339	0.266
Spatial span reversed	0.655	0.262	-0.237	0.315
Visual Patterns test span	0.625	0.319	-0.367	0.131
SWM strategy score [b]	-0.676	-0.113	0.214	0.388
SRec Correct	0.615	-0.023	0.015	-0.299
SOPT total errors [b]	-0.729	0.096	0.136	0.287
DSST	0.608	0.120	0.159	0.485
Rey total A1-5	0.690	-0.500	0.164	0.250
Rey A7	0.648	-0.618	0.110	0.117
Rey Recognition A	0.581	-0.592	0.144	-0.264
Digit Span (reverse)	0.453	0.514	0.227	-0.341
FAS correct	0.505	0.339	0.599	-0.050
ELFT correct	0.596	0.239	0.588	0.117

a. 4 components extracted.
b. Note: although loadings are negative, these variables report error scores and therefore should be reversed for interpretation on component loading.

In terms of which loadings should used for interpretation, Stevens recommends using a fixed criteria based on the sample size rather than a simple convention (such as >0.4). As the loading is effectively a Pearson correlation coefficient between each variable and the factor (the linear combination of the variables), the cut-off adopted should be the coefficient at a given significance level for the particular sample size. Therefore with n=100 and $p<0.01$, the critical value for a 2-tailed correlation coefficient is >2*(0.256)= 0.512

Table 4-5. PCA #01 Varimax rotated component matrix (initial model) a

	Component			
	1	2	3	4
SWM between errors [b]	-0.547	-0.259	-0.616	-0.021
Spatial span	0.733	0.178	0.258	0.049
Spatial span reversed	0.754	0.100	0.199	0.189
Visual Patterns test span	0.706	-0.021	0.370	0.095
SWM strategy score [b]	-0.286	-0.127	-0.742	-0.135
SRec Correct	0.153	0.275	0.560	0.234
SOPT total errors [b]	-0.273	-0.359	-0.650	-0.124
DSST	0.598	0.311	-0.088	0.427
Rey total A1-5	0.303	0.828	0.084	0.178
Rey A7	0.187	0.871	0.171	0.062
Rey Recognition A	-0.086	0.764	0.427	0.067
Digit Span (reverse)	0.110	-0.206	0.491	0.584
FAS correct	0.083	0.094	0.168	0.829
ELFT correct	0.208	0.246	0.080	0.814

Rotation Method: varimax with Kaiser Normalization. Shaded values are below the recommended critical value of 0.512

a. Rotation converged in 7 iterations. **b**. Note: although variable loadings on each component are negative, these variables report error scores and therefore should be reversed for interpretation of true component loading.

From Table 4-5 above there are four components after varimax factor rotation. The clustering of variables suggests that component 2 represents *verbal learning and memory* and component 4 *(verbal) executive function and working memory*. The remaining two components appear to represent differing aspects of visuo-spatial processing: component 1 a *short-term visuo-spatial memory* measure and component 3 a *self-ordered/strategic visuo-spatial processing* measure.

Of great interest is the fact that SWM between search errors is the only variable to load significantly across these two measures. Of the remaining measures, although not significant, it is of note that the DSST loads moderately onto components 2 and 4, possibly representing the contribution of general psychomotor speed in many different components. Similarly, reverse digit span also loads moderately onto the strategic component 3, possibly reflecting the WM manipulation component to the test.

4.3.1.2 Optimised model

Resulting from the discussion previously on the differences between orthogonal and oblique factor rotation, the direct Oblimin pattern and structure matrices were also generated and compared to the rotated model in section 4.3.1. These also provide additional information on the variables which may load onto multiple components, allowing generation of 'cleaner' composite scores. From the component correlation matrix there does appear to be some degree of correlation between the factors. This is more clearly evident when comparing the loadings of individual variables.

Table 4-6. PCA #01 component correlation matrix (optimised model) a

Component	1	2	3	4
1	1.000	-0.354	0.378	-0.230
2	-0.354	1.000	-0.243	0.151
3	0.378	-0.243	1.000	-0.164
4	-0.230	0.151	-0.164	1.000

Rotation Method: Oblimin with Kaiser Normalization.

Table 4-7. PCA #01 Oblimin rotated pattern and structure matrices (optimised model)

Oblimin Pattern Matrix	Component			
	1	2	3	4
SWM between errors [b]	-0.607	0.216	0.075	0.389
Spatial span	0.794	-0.074	-0.073	0.000
Spatial span reversed	0.798	0.028	0.080	0.068
Visual Patterns test span	0.795	0.130	-0.002	-0.146
SWM strategy score [b]	-0.324	0.118	-0.100	0.593
SRec Correct	0.119	-0.279	0.215	-0.413
SOPT total errors [b]	-0.267	0.362	-0.072	0.467
DSST	0.521	-0.186	0.342	0.374
Rey total A1-5	0.172	-0.816	0.083	0.181
Rey A7	0.066	-0.898	-0.025	0.054
Rey Recognition A	-0.203	-0.845	0.035	-0.289
Digit Span (reverse)	0.073	0.258	0.634	-0.395
FAS correct	-0.079	-0.026	0.876	0.000
ELFT correct	0.032	-0.162	0.827	0.146

Oblimin Structure Matrix	Component			
	1	2	3	4
SWM between errors [b]	-0.744	0.471	-0.271	0.549
Spatial span	0.793	-0.338	0.246	-0.183
Spatial span reversed	0.802	-0.263	0.363	-0.125
Visual Patterns test span	0.782	-0.173	0.291	-0.309
SWM strategy score [b]	-0.540	0.346	-0.349	0.702
SRec Correct	0.394	-0.436	0.396	-0.518
SOPT total errors [b]	-0.530	0.544	-0.338	0.595
DSST	0.630	-0.397	0.523	0.169
Rey total A1-5	0.451	-0.870	0.317	0.004
Rey A7	0.361	-0.907	0.209	-0.092
Rey Recognition A	0.176	-0.825	0.211	-0.375
Digit Span (reverse)	0.313	0.018	0.664	-0.477
FAS correct	0.262	-0.211	0.852	-0.130
ELFT correct	0.368	-0.352	0.854	-0.022

From the Oblique rotation reported above, it is clear that three of the components are identical to the orthogonal solution [19]. The pattern matrix shows the cluster of factor loadings in components 1, 2 and 3 are identical to components 1, 2 and 4 respectively of the varimax solution. The fourth component in the structure matrix also shows identical loadings to the varimax solution, although as can be seen from the pattern matrix, these load less cleanly due to moderate loadings with other factors. Loadings in the pattern matrix further show that Spatial Recognition and SOPT do not load uniquely onto any of the four components for the same reason.

Therefore, this model was revisited after excluding variables that were not loading sufficiently uniquely onto individual components. Exclusion was done one variable at a time until a suitable, unique factor rotation was obtained within a PCA that met all criteria described previously. After the exclusion of SOPT, DSST and SRec, the final model was achieved. The factorability of the variables was again confirmed using the criteria described above (section 4.2.3). The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.802 (the cut-off for a 'very good' value is above 0.8) and Bartlett's test of sphericity was significant (χ^2=476.9, df=55, p<0.0001). The diagonals of the anti-image correlation matrix were all much greater than 0.5 (the lowest value was 0.719), justifying the inclusion of each item in the analysis. The determinant of the initial correlation matrix was |R|=0.006 (well above the recommended 0.00001) suggesting that multicolinearity is not an issue with these data. Finally, the communalities for the PCA ranged from 0.536 to 0.772, with a mean of 0.675 (see Table 4-8 below).

[19] For clarity, cells are coloured red if the factor loadings are no longer significant, while those in green are loadings that have entered the component.

Table 4-8. PCA #01 communalities (optimised model)

	Initial	Extraction
SWM between errors	1.000	0.744
Spatial span	1.000	0.611
Spatial span reversed	1.000	0.571
Visual Patterns test span	1.000	0.622
SWM strategy score	1.000	0.536
Rey total A1-5	1.000	0.762
Rey A7	1.000	0.836
Rey Recognition A	1.000	0.689
Digit Span (reverse)	1.000	0.542
FAS correct	1.000	0.772
ELFT correct	1.000	0.744

Three components were now extracted, with each independently explaining 40.7%, 14.8%, and 12.0% of the variance (cumulatively: 40.7, 55.5, and 67.5%).

Table 4-9. PCA #01 unrotated component matrix (optimised model)

	Component		
	1	2	3
SWM between errors	-0.788	0.040	0.349
Spatial span	0.684	-0.134	-0.354
Spatial span reversed	0.661	-0.258	-0.259
Visual Patterns test span	0.620	-0.329	-0.360
SWM strategy score	-0.701	0.142	0.157
Rey total A1-5	0.690	0.523	0.112
Rey A7	0.660	0.628	0.079
Rey Recognition A	0.571	0.584	0.147
Digit Span (reverse)	0.435	-0.519	0.288
FAS correct	0.523	-0.316	0.631
ELFT correct	0.611	-0.197	0.576

After direct Oblimin rotation, the factor loadings of the pattern and structure matrices matched (in terms of individual variables loading to a component), indicating orthogonal factors.

Table 4-10. PCA #01 Oblimin rotated pattern and structure matrices (optimised model)

Pattern Matrix

	Component		
	1	2	3
SWM between errors [b]	-0.797	-0.213	0.072
Spatial span	0.776	0.080	-0.065
Spatial span reversed	0.731	-0.024	0.074
Visual Patterns test span	0.821	-0.132	0.000
SWM strategy score [b]	-0.622	-0.126	-0.128
Rey total A1-5	0.116	0.809	0.073
Rey A7	0.084	0.887	-0.015
Rey Recognition A	-0.011	0.825	0.042
Digit Span (reverse)	0.233	-0.236	0.639
FAS correct	-0.086	0.073	0.894
ELFT correct	-0.034	0.210	0.813

Structure Matrix

	Component		
	1	2	3
SWM between errors [b]	-0.837	-0.452	-0.273
Spatial span	0.776	0.313	0.247
Spatial span reversed	0.752	0.222	0.350
Visual Patterns test span	0.779	0.128	0.290
SWM strategy score [b]	-0.711	-0.346	-0.390
Rey total A1-5	0.401	0.860	0.267
Rey A7	0.360	0.911	0.180
Rey Recognition A	0.266	0.829	0.189
Digit Span (reverse)	0.403	-0.045	0.685
FAS correct	0.280	0.210	0.874
ELFT correct	0.344	0.348	0.839

b. Note: although variable loadings on each component are negative, these variables report error scores and therefore should be reversed for interpretation of true component loading.

To confirm this, an orthogonal varimax rotation was also performed and produced the expected identical cluster of loadings.

Table 4-11. PCA #01 Varimax rotated component matrix (optimised model)a

	Component		
	1	2	3
SWM between errors [b]	-0.788	-0.337	-0.090
Spatial span	0.750	0.203	0.085
Spatial span reversed	0.719	0.107	0.206
Visual Patterns test span	0.776	0.007	0.143
SWM strategy score [b]	-0.645	-0.241	-0.248
Rey total A1-5	0.243	0.826	0.142
Rey A7	0.206	0.889	0.054
Rey Recognition A	0.116	0.817	0.089
Digit Span (reverse)	0.315	-0.131	0.652
FAS correct	0.099	0.145	0.861
ELFT correct	0.154	0.281	0.801

Rotation Method: varimax with Kaiser Normalization. **a**. Rotation converged in 5 iterations. **b**. Note: although variable loadings on each component are negative, these variables report error scores and therefore should be reversed for interpretation of true component loading.

From these loadings the three components are clearly evident as: visuo-spatial memory in component 1, *verbal learning and memory* in component 2, and *(verbal) executive function and working memory* in component 3. These latter two are identical to the initial models, however through the removal of tests that exhibit a broader profile of loadings, the visuo-spatial components have collapsed into a single factor.

4.3.2 PCA #02: Total Sample (for SWM sub-analysis)

4.3.2.1 Initial model

As discussed earlier, one of the aims of the subsequent chapters is to reduce the neuropsychological variables into composites that can be used to examine the relationship with specific spatial tests and processes. To this end, the PCA was re-run after excluding SWM entirely in order to use this test as the dependent variable.

Interestingly, the pattern of excluded variables followed closely that of PCA #01 in that PRec-m, SRec and also DSST did not load significantly after factor rotation (and also exhibited low communality) and were therefore excluded. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.784 (the cut-off for a 'good' value is between 0.7 and 0.8) and Bartlett's test of sphericity was significant (χ^2=398.5, df=45, p<0.0001). The diagonals of the anti-image correlation matrix were all much greater than 0.5 (the lowest value was 0.696), justifying the inclusion of each item in the analysis. The determinant of the initial correlation matrix was |R|=0.015 (well above the recommended 0.00001) suggesting that multicolinearity is not an issue with these data. Finally, the communalities for the PCA ranged from 0.546 to 0.839, with a mean of 0.690 (see Table 4-12 below).

Table 4-12. PCA #02 communalities (initial model)

	Initial	Extraction
Spatial span	1.000	0.649
Spatial span reversed	1.000	0.668
Visual Patterns test span	1.000	0.659
SOPT total errors	1.000	0.574
Rey total A1-5	1.000	0.760
Rey A7	1.000	0.839
Rey Recognition A	1.000	0.706
Digit span (reverse)	1.000	0.546
FAS correct	1.000	0.768
ELFT correct	1.000	0.727

Three components were extracted explaining 40.5%, 16.2% and 12.3% of the variance (40.5, 56.7, 69.0% respectively).

Table 4-13. PCA #02 unrotated component matrix (initial model) a

	Component		
	1	2	3
Spatial span	0.653	0.153	-0.445
Spatial span reversed	0.653	0.296	-0.392
Visual Patterns test span	0.591	0.333	-0.446
SOPT total errors	-0.743	0.051	0.141
Rey total A1-5	0.738	-0.460	0.055
Rey A7	0.707	-0.581	0.051
Rey Recognition A	0.599	-0.558	0.189
Digit span (reverse)	0.439	0.549	0.229
FAS correct	0.554	0.394	0.552
ELFT correct	0.623	0.278	0.511

a. 3 components extracted.

After varimax rotation, the same 3 components were evident: a verbal learning and memory component (1); a visuo-spatial memory component (2); and a verbal executive/WM component (3).

Table 4-14. PCA #02 Varimax rotated component matrix (initial model) a

	Component		
	1	2	3
Spatial span	0.228	0.768	0.085
Spatial span reversed	0.130	0.782	0.199
Visual Patterns test span	0.054	0.795	0.154
SOPT total errors	-0.490	-0.534	-0.223
Rey total A1-5	0.827	0.244	0.131
Rey A7	0.896	0.184	0.047
Rey Recognition A	0.833	0.030	0.104
Digit span (reverse)	-0.094	0.305	0.667
FAS correct	0.149	0.091	0.859
ELFT correct	0.273	0.120	0.799

Rotation Method: varimax with Kaiser Normalization. a. Rotation converged in 5 iterations.

4.3.2.2 Optimised model

Again, as before, although not significant above the defined cut-off for interpretation it is evident that SOPT also loads onto component 1. Therefore when comparing this to the oblique rotation method, SOPT did not load uniquely onto any component. Once removed, the Oblimin method produced an identical pattern and structure matrix which was confirmed through orthogonal rotation.

This new PCA contained 9 variables. The KMO statistic was 0.758 and Bartlett's test was significant (χ^2=334.0, df=36, p<0.0001). The diagonals of the anti-image correlation matrix were all much greater than 0.5 (the lowest value was 0.682), justifying the inclusion of each item in the analysis. The determinant of the initial correlation matrix was |R|=0.030 (well above the recommended 0.00001). Finally, the communalities for the PCA ranged from 0.547 to 0.848, with a mean of 0.711 (see Table 4-15).

Table 4-15. PCA #02 communalities (optimised model)

	Initial	Extraction
Spatial span	1.000	0.688
Spatial span reversed	1.000	0.706
Visual Patterns test span	1.000	0.641
Rey total A1-5	1.000	0.786
Rey A7	1.000	0.848
Rey Recognition A	1.000	0.691
Digit span (reverse)	1.000	0.547
FAS correct	1.000	0.763
ELFT correct	1.000	0.731

Three components were extracted explaining 39.6%, 18.0% and 13.5% of the variance (cumulatively; 39.6, 57.6, 71.1%).

Table 4-16. PCA #02 unrotated component matrix (optimised model) a

	Component		
	1	2	3
Spatial span	0.660	0.146	-0.481
Spatial span reversed	0.664	0.287	-0.428
Visual Patterns test span	0.574	0.333	-0.449
Rey total A1-5	0.749	-0.475	0.011
Rey A7	0.705	-0.593	0.018
Rey Recognition A	0.582	-0.565	0.182
Digit span (reverse)	0.429	0.546	0.256
FAS correct	0.581	0.377	0.532
ELFT correct	0.667	0.257	0.469

a. 3 components extracted.

After direct Oblimin rotation, 3 components were evident in the pattern and structure matrices: 1) *a visuo-spatial memory* component; 2) a *verbal learning and memory* component; and 3) a *verbal executive/WM* component.

Table 4-17. PCA #02 Oblimin rotated pattern and structure matrices (optimised model) a

Pattern Matrix [a]

	Component		
	1	2	3
Spatial span	0.804	-0.141	-0.062
Spatial span reversed	0.809	-0.028	0.062
Visual Patterns test span	0.802	0.064	0.034
Rey total A1-5	0.168	-0.818	0.038
Rey A7	0.098	-0.899	-0.039
Rey Recognition A	-0.099	-0.838	0.065
Digit span (reverse)	0.166	0.206	0.685
FAS correct	-0.077	-0.071	0.882
ELFT correct	-0.023	-0.212	0.797

Rotation Method: Oblimin with Kaiser Normalization. a. Rotation converged in 7 iterations.

Structure Matrix

	Component		
	1	2	3
Spatial span	0.817	-0.333	0.252
Spatial span reversed	0.838	-0.246	0.355
Visual Patterns test span	0.798	-0.147	0.307
Rey total A1-5	0.390	-0.869	0.261
Rey A7	0.314	-0.917	0.175
Rey Recognition A	0.137	-0.826	0.197
Digit span (reverse)	0.358	0.028	0.703
FAS correct	0.255	-0.227	0.869
ELFT correct	0.314	-0.364	0.830

Rotation Method: Oblimin with Kaiser Normalization.

As can be seen from the convergence of the pattern and structure matrices, the factors are now orthogonal and are identical to the varimax rotated solution.

Table 4-18. PCA #02 Varimax rotated component matrix (optimised model) a

	Component		
	1	2	3
Spatial span	0.240	0.789	0.092
Spatial span reversed	0.141	0.803	0.205
Visual Patterns test span	0.047	0.781	0.170
Rey total A1-5	0.837	0.262	0.131
Rey A7	0.900	0.190	0.049
Rey Recognition A	0.824	0.008	0.112
Digit span (reverse)	-0.111	0.258	0.684
FAS correct	0.152	0.086	0.856
ELFT correct	0.290	0.140	0.792

Rotation Method: varimax with Kaiser Normalization. a. Rotation converged in 5 iterations.

4.3.3 PCA #03: Control participants

4.3.3.1 Initial model

In order to examine any differences in the structure of the models between patients and controls, the PCA was re-run for each group separately. To allow for differences to emerge, the analyses were performed from the point of initial data screening. Examination of the correlation matrix led to the exclusion of the digit span forward and both SCOLP measures. From this analysis it emerged that SWM within search errors and SRec did not load significantly into the rotated solution and the modified PREC exhibited low communality; all were excluded.

This final model contained 13 variables. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.738 and Bartlett's test of sphericity was significant (χ^2=281.9, df=78, p<0.0001). The diagonals of the anti-image correlation matrix were all greater than 0.5 justifying the inclusion of each item in the analysis. The determinant of the initial correlation matrix was |R|=0.001 suggesting that multicolinearity is not an issue with these data. Finally, the communalities for the PCA ranged from 0.610 to 0.870, with a mean of 0.733 (see Table 4-19 below).

Table 4-19. PCA #03 communalities (initial model)

	Initial	Extraction
SWM between errors	1.000	0.720
Spatial span	1.000	0.694
Spatial span reversed	1.000	0.621
Visual Patterns test span	1.000	0.700
SWM strategy score	1.000	0.756
SOPT total errors	1.000	0.758
DSST	1.000	0.610
Rey total A1-5	1.000	0.823
Rey A7	1.000	0.870
Rey Recognition A	1.000	0.751
Digit Span (reverse)	1.000	0.675
FAS correct	1.000	0.766
ELFT correct	1.000	0.789

Four components were extracted explaining 35.6%, 15.1%, 13.1%, 9.5% (cumulatively, 35.6, 50.7, 63.8, 73.3%) of the variance.

Table 4-20. PCA #03 unrotated component matrix (initial model) a

	Component			
	1	2	3	4
SWM between errors	-0.778	0.139	0.303	0.056
Spatial span	0.687	-0.259	-0.218	0.327
Spatial span reversed	0.568	0.035	-0.185	0.513
Visual Patterns test span	0.547	-0.117	-0.611	0.117
SWM strategy score	-0.661	-0.039	0.365	0.429
SOPT total errors	-0.749	-0.002	0.093	0.433
DSST	0.521	0.175	0.088	0.549
Rey total A1-5	0.608	-0.284	0.601	0.103
Rey A7	0.695	-0.252	0.570	-0.008
Rey Recognition A	0.627	-0.267	0.452	-0.288
Digit Span (reverse)	0.521	0.574	-0.131	-0.239
FAS correct	0.236	0.805	0.242	-0.059
ELFT correct	0.298	0.799	0.189	0.163

a. 4 components extracted.

After varimax rotation, the criteria for significance of the loadings were again calculated for the obtained sample size. At an alpha level of $p<0.05$ and $n=47$, the cut-off for interpretation was $>2*(0.2876)= 0.575$ [20]

[20] Note that this is more stringent than the cut-off with the overall sample, hence the use of $p<0.05$ rather than 0.01.

Table 4-21. PCA #03 Varimax rotated component matrix (initial model) a

	Component			
	1	2	3	4
SWM between errors [b]	0.677	-0.240	-0.451	0.022
Spatial span	-0.355	0.254	0.697	-0.135
Spatial span reversed	-0.151	0.096	0.757	0.123
Visual Patterns test span	-0.595	-0.146	0.554	-0.131
SWM strategy score [b]	0.851	-0.110	-0.095	-0.101
SOPT total errors [b]	0.765	-0.380	-0.065	-0.154
DSST	0.045	0.229	0.678	0.310
Rey total A1-5	-0.005	0.871	0.251	0.025
Rey A7	-0.139	0.895	0.212	0.067
Rey Recognition A	-0.322	0.805	-0.015	0.008
Digit Span (reverse)	-0.522	0.015	0.079	0.629
FAS correct	-0.041	0.048	-0.040	0.872
ELFT correct	0.022	0.019	0.186	0.868

a. Rotation converged in 6 iterations. [b] n.b. for these measures, higher scores equate to worse performance.

After rotation, there are four components that emerge. Component 2 and 4 are identical to those seen in the overall group analysis, and represent verbal learning and memory and (verbal) executive function/WM respectively. For component 1, it appear that in controls this represents more generically an executive function/WM factor which seems to be domain non-specific, including both verbal and visuo-spatial elements. Component 3 appears to relate more closely to the visuo-spatial sketchpad and executive control of this slave system, as tests that assess the processing of the inner eye (VPT) and inner scribe (SSP) are included.

However, it is also noted that a number of measures load across multiple components, therefore an optimised model was sought to produce more independent factors.

4.3.3.2 Optimised model

Using an oblique rotation method it was noted that the VPT loaded heavily onto 2 components; removal of this factor produced a model in which reverse digit span did not load uniquely in the pattern matrix (due to shared variance with other components). Removal of this led to the final model below.

The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.705 and Bartlett's test of sphericity was significant (χ^2=228.7, df=55, p<0.0001). The diagonals of the anti-image correlation matrix were all greater than 0.5 justifying the inclusion of each item in the analysis. The determinant of the initial correlation matrix was |R|=0.004 suggesting that multicolinearity is not an issue with these data. Finally, the communalities for the PCA ranged from 0.605 to 0.882, with a mean of 0.764 (see Table 4-22 below).

Table 4-22. PCA #03 communalities (optimised model)

	Initial	Extraction
SWM between errors	1.000	0.745
Spatial span	1.000	0.678
Spatial span (reverse)	1.000	0.690
SWM strategy score	1.000	0.805
DSST	1.000	0.605
SOPT total errors	1.000	0.755
Rey total A1-5	1.000	0.823
Rey A7	1.000	0.882
Rey Recognition A	1.000	0.755
FAS correct	1.000	0.827
ELFT correct	1.000	0.838

Extraction Method: Principal Component Analysis.

Table 4-23. PCA #03 unrotated component matrix (optimised model) a

	Component			
	1	2	3	4
SWM between errors	-0.748	0.135	0.403	0.063
Spatial span	0.678	-0.172	-0.267	0.342
Spatial span (reverse)	0.549	0.110	-0.328	0.518
SWM strategy score	-0.607	0.060	0.483	0.446
DSST	0.519	0.317	-0.018	0.486
SOPT total errors	-0.722	0.048	0.170	0.450
Rey total A1-5	0.701	-0.126	0.547	0.132
Rey A7	0.772	-0.136	0.516	0.034
Rey Recognition A	0.697	-0.197	0.399	-0.268
FAS correct	0.197	0.844	0.118	-0.250
ELFT correct	0.261	0.877	0.025	-0.033

Extraction Method: Principal Component Analysis. a. 4 components extracted.

From this model, 4 components were again extracted explaining 37.8%, 15.6%, 12.2% and 10.7% (37.8, 53.5, 65.7, and 76.4% cumulatively).

Table 4-24. PCA #03 Oblimin rotated pattern and structure matrices (optimised model)

Pattern Matrix [a]

	Component			
	1	2	3	4
SWM between errors	-0.665	0.080	-0.052	-0.377
Spatial span	0.261	-0.202	0.132	0.653
Spatial span (reverse)	0.087	0.009	-0.067	0.822
SWM strategy score	-0.917	-0.057	0.097	-0.022
SOPT total errors	-0.758	-0.117	-0.252	0.070
DSST	-0.137	0.249	0.157	0.675
Rey total A1-5	-0.136	-0.014	0.900	0.131
Rey A7	-0.014	0.004	0.915	0.082
Rey Recognition A	0.251	-0.007	0.796	-0.175
FAS correct	0.083	0.910	0.022	-0.103
ELFT correct	0.018	0.888	-0.049	0.159

Structure Matrix

	Component			
	1	2	3	4
SWM between errors	-0.776	0.015	-0.384	-0.552
Spatial span	0.466	-0.111	0.401	0.736
Spatial span (reverse)	0.274	0.099	0.211	0.825
SWM strategy score	-0.891	-0.075	-0.224	-0.232
SOPT total errors	-0.828	-0.150	-0.496	-0.211
DSST	0.093	0.334	0.334	0.715
Rey total A1-5	0.201	0.071	0.893	0.366
Rey A7	0.317	0.088	0.936	0.355
Rey Recognition A	0.476	0.045	0.827	0.128
FAS correct	0.088	0.902	0.093	0.028
ELFT correct	0.065	0.902	0.077	0.250

a. Rotation converged in 8 iterations.

Again, the factor loadings above were compared to the varimax rotated model and an identical pattern of loading was observed between this and the Oblimin pattern matrix.

Table 4-25. PCA #03 Varimax rotated component matrix (optimised model)

	Component			
	1	2	3	4
SWM between errors	-0.193	0.707	-0.454	0.048
Spatial span	0.240	-0.357	0.685	-0.154
Spatial span (reverse)	0.049	-0.173	0.809	0.059
SWM strategy score	-0.047	0.886	-0.117	-0.063
SOPT total errors	-0.352	0.781	-0.067	-0.129
DSST	0.223	0.024	0.684	0.295
Rey total A1-5	0.868	-0.042	0.258	0.029
Rey A7	0.896	-0.159	0.228	0.045
Rey Recognition A	0.787	-0.367	-0.014	0.016
FAS correct	0.052	-0.079	-0.045	0.903
ELFT correct	0.005	-0.034	0.193	0.895

Rotation Method: Varimax with Kaiser Normalization. a. Rotation converged in 5 iterations.

This four component solution retains the verbal memory and learning factor in component 1, and component 2 appears to be a strategic, visuo-spatial self-ordered search component, while 3 includes the immediate spatial span measures with psychomotor speed. In component 4 the digit span was not included in the model leaving a verbal fluency/executive component

4.3.4 PCA #04: Control participants (for SWM sub-analysis)

4.3.4.1 Initial model

Again, a series of components was sought that could be used to produce a composite score that could be used to predict SWM performance. After removal of the SWM variables, the initial analysis revealed that SRec did not load onto any component and was removed to produce the model below. All criteria were met as before regarding a significant Bartlett's test (χ^2=223.0, df=66, p<0.0001), and acceptable KMO of 0.691 and the diagonals of the anti-image correlation matrix were all greater than 0.5. The determinant of the initial correlation matrix was |R|=0.004 suggesting that multicolinearity is not an issue with these data. Finally, the communalities for the PCA ranged from 0.487 to 0.880, with a mean of 0.716 (see Table 4-26 below).

Table 4-26. PCA #04 communalities (initial model)

	Initial	Extraction
PRec modified	1.000	0.487
Visual Patterns test	1.000	0.716
Spatial span	1.000	0.750
Spatial span reversed	1.000	0.624
Rey total A1-5	1.000	0.813
Rey A7	1.000	0.880
Rey Recognition A	1.000	0.765
Digit span (reverse)	1.000	0.583
FAS correct	1.000	0.775
ELFT correct	1.000	0.782
DSST	1.000	0.631
SOPT v2 total errors	1.000	0.783

From the four components extracted, the proportion of variance explained was 32.7, 16.2, 13.0, and 9.6% (32.7, 48.9, 61.9, and 71.6% respectively).

Table 4-27. PCA #04 unrotated component matrix (initial model) a

	Component			
	1	2	3	4
PRec modified	0.414	0.054	-0.253	-0.498
Visual Patterns test	0.474	-0.073	-0.694	-0.058
Spatial span	0.681	-0.264	-0.414	0.209
Spatial span reversed	0.587	0.037	-0.385	0.360
Rey total A1-5	0.665	-0.354	0.470	0.155
Rey A7	0.749	-0.330	0.451	0.080
Rey Recognition A	0.656	-0.323	0.396	-0.269
Digit span (reverse)	0.493	0.560	-0.047	-0.154
FAS correct	0.313	0.775	0.252	-0.116
ELFT correct	0.358	0.768	0.185	0.172
DSST	0.533	0.164	-0.015	0.565
SOPT total errors	-0.726	0.021	0.112	0.492

a. 4 components extracted.

After varimax rotation, there were four distinct factors extracted.

Table 4-28. PCA #04 Varimax rotated component matrix (initial model) a

	Component			
	1	2	3	4
PRec modified	0.065	0.105	0.087	0.681
Visual Patterns test	-0.122	-0.083	0.647	0.525
Spatial span	0.278	-0.105	0.761	0.287
Spatial span reversed	0.114	0.150	0.759	0.109
Rey total A1-5	0.879	0.030	0.198	-0.025
Rey A7	0.908	0.071	0.211	0.079
Rey Recognition A	0.803	0.019	-0.011	0.345
Digit span (reverse)	0.044	0.656	0.174	0.347
FAS correct	0.035	0.866	-0.096	0.118
ELFT correct	0.029	0.870	0.136	-0.070
DSST	0.271	0.349	0.619	-0.229
SOPT total errors	-0.385	-0.179	-0.183	-0.754

Rotation Method: varimax with Kaiser Normalization. a. Rotation converged in 6 iterations.

Of interest is that again the same two initial components, verbal learning and memory, and verbal executive function/WM emerge, as they have done in the overall sample. The remaining two components are close to the initial model prior to removing SWM, with component 3 being an executive/spatial sketchpad component, and 4 being a visual memory component.

4.3.4.2 Optimised model

As with previous models, the PCA was also examined using an Oblimin rotation method. Using the same variables as in the initial model indicated that DSST and SOPT did not produce significant unique loadings onto a single component (from the pattern matrix) and were

therefore excluded. PRec-m exhibited a very low communality (0.244) and was also excluded leaving 9 variables for inclusion.

Criteria were met as before regarding a significant Bartlett's test (χ^2=161.2, df=36, p<0.0001), and acceptable KMO of 0.654 and the diagonals of the anti-image correlation matrix were all greater than 0.5. The determinant of the initial correlation matrix was |R|=0.022. Finally, the communalities for the PCA ranged from 0.565 to 0.884, with a mean of 0.723 (see Table 4-29 below).

Table 4-29. PCA #04 communalities (optimised model)

	Initial	Extraction
Visual Patterns Test	1.000	0.663
Spatial span	1.000	0.777
Spatial span (reverse)	1.000	0.604
Rey total A1-5	1.000	0.807
Rey A7	1.000	0.884
Rey Recognition A	1.000	0.680
Digit Span (reverse)	1.000	0.565
FAS correct	1.000	0.775
ELFT correct	1.000	0.753

From the PCA, three components were extracted explaining 34.1%, 21.4%, and 16.8% of the variance (cumulatively 34.1%, 55.5%, 72.3%).

Table 4-30. PCA #04 unrotated component matrix (optimised model)

	Component		
	1	2	3
Visual Patterns Test	0.400	-0.080	0.705
Spatial span	0.687	-0.237	0.499
Spatial span (reverse)	0.599	0.058	0.492
Rey total A1-5	0.749	-0.288	-0.404
Rey A7	0.826	-0.252	-0.372
Rey Recognition A	0.669	-0.260	-0.406
Digit Span (reverse)	0.438	0.600	0.113
FAS correct	0.277	0.811	-0.203
ELFT correct	0.340	0.793	-0.092

a. 3 components extracted.

From the pattern and structure matrices produced after Oblimin rotation it can be seen that the two converge with the same pattern of loadings, again indicating an orthogonal structure which can be confirmed by entering the variable into a varimax rotated model.

The three components represent a verbal learning and memory component, a verbal executive/WM component and a visuo-spatial memory (or more specifically the visuo-spatial sketchpad and executive control). It is also clear from these components that the variables included in each one load very highly onto those components, with negligible loadings onto the others.

Table 4-31. PCA #04 Oblimin rotated pattern and structure matrices (optimised model)

Pattern Matrix [a]

	Component		
	1	2	3
Visual Patterns Test	-0.156	-0.050	0.842
Spatial span	0.240	-0.095	0.801
Spatial span (reverse)	0.071	0.165	0.718
Rey total A1-5	0.899	-0.014	0.002
Rey A7	0.918	0.038	0.062
Rey Recognition A	0.834	-0.010	-0.040
Digit Span (reverse)	0.000	0.688	0.230
FAS correct	0.016	0.884	-0.161
ELFT correct	-0.006	0.871	-0.027

a. Rotation converged in 5 iterations.

Structure Matrix

	Component		
	1	2	3
Visual Patterns Test	0.041	0.035	0.798
Spatial span	0.423	0.030	0.847
Spatial span (reverse)	0.262	0.260	0.755
Rey total A1-5	0.898	0.086	0.217
Rey A7	0.937	0.148	0.288
Rey Recognition A	0.823	0.079	0.160
Digit Span (reverse)	0.133	0.716	0.314
FAS correct	0.077	0.867	-0.049
ELFT correct	0.085	0.867	0.077

Rotation Method: Oblimin with Kaiser Normalization.

Table 4-32. PCA #04 Varimax rotated component matrix (optimised model)

	Component		
	1	2	3
Visual Patterns Test	-0.059	-0.011	0.812
Spatial span	0.327	-0.036	0.818
Spatial span (reverse)	0.162	0.209	0.731
Rey total A1-5	0.891	0.034	0.112
Rey A7	0.919	0.090	0.177
Rey Recognition A	0.821	0.033	0.063
Digit Span (reverse)	0.062	0.699	0.269
FAS correct	0.043	0.873	-0.104
ELFT correct	0.036	0.867	0.025

Rotation Method: varimax with Kaiser Normalization. a. Rotation converged in 4 iterations.

4.3.5 PCA #05: Bipolar patients

For the PCA including patients only, the correlation matrix was used to screen data for any variables that did not correlate strongly or consistently with other variables. At this stage, SWM within errors, PRec-m, Digit span (forward), and the SCOLP measures were excluded.

4.3.5.1 Initial model

This first model showed that the ELFT did not load significantly onto any component and was removed, then the DSST was removed from the subsequent model for the same reason leaving 12 variables in the final model. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.773 and Bartlett's test of sphericity was significant (χ^2=264.6, df=66, p<0.0001). The diagonals of the anti-image correlation matrix were all greater than 0.5 justifying the inclusion of each item in the analysis. The determinant of the initial correlation matrix was |R|=0.004 suggesting that multicolinearity is not an issue with these data. Finally, the communalities for the PCA ranged from 0.442 to 0.811, with a mean of 0.645 (see Table 4-33).

Table 4-33. PCA #05 communalities (initial model)

	Initial	Extraction
SWM between search errors	1.000	0.765
Spatial span	1.000	0.699
Spatial span (reverse)	1.000	0.604
Visual Patterns test span	1.000	0.537
SWM strategy score	1.000	0.607
SRec Correct	1.000	0.694
SOPT total errors	1.000	0.597
Rey total A1-5	1.000	0.680
Rey A7	1.000	0.811
Rey Recognition A	1.000	0.680
Digit span (reverse)	1.000	0.626
FAS correct	1.000	0.442

Table 4-34. PCA #05 unrotated component matrix (initial model)

	Component		
	1	2	3
SWM between search errors	-0.839	-0.007	0.248
Spatial span	0.642	0.291	-0.450
Spatial span (reverse)	0.691	0.354	0.021
Visual Patterns test span	0.647	0.333	0.083
SWM strategy score	-0.711	-0.064	0.311
SRec Correct	0.502	0.161	0.645
SOPT total errors	-0.732	0.141	-0.203
Rey total A1-5	0.652	-0.497	0.091
Rey A7	0.544	-0.715	-0.064
Rey Recognition A	0.509	-0.589	0.272
Digit span (reverse)	0.284	0.595	0.437
FAS correct	0.561	0.299	-0.196

a. 3 components extracted.

Three components were extracted, explaining 39.0%, 16.0% and 9.5% of the variance (cumulatively, 39.0, 55.0, 64.5%).

After varimax rotation, the loadings onto each factor suggested that there was a much broader loading onto the first factor which covered executive control (as well as strategic aspects) and visuo-spatial memory. In component 2 the verbal learning and memory measures were included along with SOPT, possibly suggesting that this test was being approached in a different way compared to controls. The final component includes SRec and a verbal WM measure. This last point is of note as all the components include a mixture of verbal and visual/spatial measures.

Table 4-35. PCA #05 Varimax rotated component matrix (initial model) a

	Component		
	1	2	3
SWM between search errors	-0.761	-0.416	-0.112
Spatial span	0.835	0.035	0.013
Spatial span (reverse)	0.632	0.113	0.438
Visual Patterns test span	0.557	0.118	0.462
SWM strategy score	-0.724	-0.284	-0.042
SRec Correct	0.070	0.293	0.777
SOPT total errors	-0.375	-0.569	-0.364
Rey total A1-5	0.245	0.784	0.076
Rey A7	0.170	0.863	-0.191
Rey Recognition A	0.004	0.815	0.125
Digit span (reverse)	0.188	-0.220	0.736
FAS correct	0.636	0.036	0.190

Rotation Method: varimax with Kaiser Normalization. a. Rotation converged in 5 iterations.

4.3.5.2 Optimised model

To optimise the model, the same series of procedures were completed as above. The ELFT, SOPT and then DSST variables were excluded due to high multiple loadings across factors. The final model contained 11 variables. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.731 and Bartlett's test of sphericity was significant (χ^2=235.4, df=55, p<0.0001). The diagonals of the anti-image correlation matrix were all greater than 0.5 justifying the inclusion of each item in the analysis. The determinant of the initial correlation matrix was |R|=0.007 suggesting that multicolinearity is not an issue with these data. Finally, the communalities for the PCA ranged from 0.436 to 0.810, with a mean of 0.657 (see Table 4-36).

Table 4-36. PCA #05 communalities (optimised model)

	Initial	Extraction
SWM between search errors	1.000	0.768
Spatial span	1.000	0.689
Spatial span (reverse)	1.000	0.600
Visual Patterns test	1.000	0.519
SWM strategy score	1.000	0.602
SRec Correct	1.000	0.760
Rey total A1-5	1.000	0.678
Rey A7	1.000	0.810
Rey Recognition A	1.000	0.725
Digit span (reverse)	1.000	0.636
FAS correct	1.000	0.436

Three components were extracted, explaining 38.2%, 17.3% and 10.1% of the variance (cumulatively, 38.2, 55.5, 65.7, and 73.8%).

Table 4-37. PCA #05 unrotated component matrix (optimised model)

Component Matrix (unrotated) [a]

	Component		
	1	2	3
SWM between search errors	-.849	.031	.215
Spatial span	.674	.250	-.415
Spatial span (reverse)	.700	.333	.004
Visual Patterns test	.644	.322	.043
SWM strategy score	-.735	-.023	.248
SRec Correct	.505	.150	.695
Rey total A1-5	.630	-.517	.116
Rey A7	.520	-.734	-.037
Rey Recognition A	.490	-.606	.343
Digit span (reverse)	.290	.596	.443
FAS correct	.584	.269	-.150

a. 3 components extracted.

After rotation, the same components were observed with the exception of the exclusion of the SOPT. As the pattern and structure matrices converged in terms of their loadings, an orthogonal structure could be assumed, which was confirmed by varimax rotation.

Table 4-38. PCA #05 Oblimin rotated pattern and structure matrices (optimised model)

Pattern Matrix [a]

	Component		
	1	2	3
SWM between search errors	-0.766	0.296	0.048
Spatial span	0.880	0.092	-0.164
Spatial span (reverse)	0.658	0.014	0.270
Visual Patterns test	0.585	0.013	0.292
SWM strategy score	-0.724	0.191	0.081
SRec Correct	-0.017	-0.305	0.817
Rey total A1-5	0.188	-0.753	0.036
Rey A7	0.117	-0.850	-0.215
Rey Recognition A	-0.103	-0.853	0.189
Digit span (reverse)	0.168	0.264	0.708
FAS correct	0.646	0.056	0.076

Rotation Method: Oblimin with Kaiser Normalization. a. Rotation converged in 7 iterations.

Structure Matrix

	Component		
	1	2	3
SWM between search errors	-0.826	0.486	-0.172
Spatial span	0.811	-0.123	0.081
Spatial span (reverse)	0.730	-0.154	0.454
Visual Patterns test	0.664	-0.138	0.456
SWM strategy score	-0.748	0.369	-0.125
SRec Correct	0.288	-0.316	0.818
Rey total A1-5	0.384	-0.800	0.103
Rey A7	0.267	-0.875	-0.166
Rey Recognition A	0.162	-0.831	0.177
Digit span (reverse)	0.301	0.209	0.750
FAS correct	0.653	-0.106	0.256

Rotation Method: Oblimin with Kaiser Normalization.

Table 4-39. PCA #05 Varimax rotated component matrix (optimised model) a

	Component		
	1	2	3
SWM between search errors	-0.774	-0.403	-0.079
Spatial span	0.829	0.033	-0.019
Spatial span (reverse)	0.673	0.080	0.376
Visual Patterns test	0.605	0.070	0.386
SWM strategy score	-0.718	-0.292	-0.039
SRec Correct	0.115	0.302	0.810
Rey total A1-5	0.270	0.775	0.070
Rey A7	0.181	0.861	-0.191
Rey Recognition A	0.016	0.833	0.175
Digit span (reverse)	0.220	-0.237	0.729
FAS correct	0.634	0.036	0.181

Rotation Method: varimax with Kaiser Normalization. a. Rotation converged in 6 iterations.

4.3.6 PCA #06: Bipolar patients (for SWM sub-analysis)

Finally, the PCA for use as a predictor of SWM performance was obtained. After screening the variables, the ELFT was excluded as it did not load significantly onto any component and then the DSST was excluded with a low communality (0.378). The pattern of loadings in the final model was identical to the overall model above (PCA #05) but with SWM omitted.

4.3.6.1 Initial analysis

In brief, all conditions as laid out were met: the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.734 and Bartlett's test of sphericity was significant (χ^2=180.9, df=45, p<0.0001). The diagonals of the anti-image correlation matrix were all greater than 0.5 justifying the inclusion of each item in the analysis. The determinant of the initial correlation

matrix was |R|=0.023 suggesting that multicolinearity is not an issue with these data. Finally, the communalities for the PCA ranged from 0.468 to 0.824, with a mean of 0.663 (see Table 4-40).

Table 4-40. PCA #06 communalities (initial model) a

	Initial	Extraction
SRec Correct	1.000	0.711
Visual Patterns test span	1.000	0.570
Spatial span	1.000	0.703
Spatial span (reverse)	1.000	0.636
Rey total A1-5	1.000	0.724
Rey A7	1.000	0.824
Rey Recognition A	1.000	0.718
Digit span (reverse)	1.000	0.664
FAS correct	1.000	0.468
SOPT total errors	1.000	0.607

Three components were extracted explaining 36.6%, 19.2% and 10.5% of the variance (cumulatively 36.6, 55.7, and 66.3%).

As discussed above, these components parallel the earlier analysis, but have the SWM loadings removed.

Table 4-41. PCA #06 unrotated component matrix (initial model) a

	Component		
	1	2	3
SRec Correct	0.557	0.180	0.607
Visual Patterns test span	0.663	0.352	-0.082
Spatial span	0.596	0.295	-0.510
Spatial span (reverse)	0.694	0.372	-0.128
Rey total A1-5	0.701	-0.478	-0.059
Rey A7	0.558	-0.704	-0.131
Rey Recognition A	0.523	-0.578	0.332
Digit span (reverse)	0.304	0.606	0.451
FAS correct	0.558	0.306	-0.250
SOPT total errors	-0.767	0.120	-0.070

a. 3 components extracted.

Table 4-42. PCA #06 Varimax rotated component matrix (initial model) a

	Component		
	1	2	3
SRec Correct	0.287	0.136	0.781
Visual Patterns test span	0.135	0.668	0.325
Spatial span	0.083	0.831	-0.077
Spatial span (reverse)	0.133	0.724	0.307
Rey total A1-5	0.800	0.290	0.000
Rey A7	0.875	0.124	-0.209
Rey Recognition A	0.814	-0.095	0.215
Digit span (reverse)	-0.219	0.248	0.745
FAS correct	0.084	0.667	0.128
SOPT total errors	-0.582	-0.433	-0.285

Rotation Method: varimax with Kaiser Normalization. a. Rotation converged in 5 iterations.

4.3.6.2 Optimised model

From the above PCA, it can be seen that the SOPT loaded moderately onto a second component. When the Oblimin rotation was performed on the variables, SOPT was excluded as is did not load independently onto any component in the pattern matrix. Again the final model is identical to the optimised overall patient model but with SWM omitted.

The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.671 and Bartlett's test of sphericity was significant (χ^2=151.2, df=36, p<0.0001). The diagonals of the anti-image correlation matrix were all greater than 0.5 justifying the inclusion of each item in the analysis. The determinant of the initial correlation matrix was |R|=0.043 suggesting that multicolinearity is not an issue with these data. Finally, the communalities for the PCA ranged from 0.476 to 0.829, with a mean of 0.678 (see Table 4-43).

Table 4-43. PCA #06 communalities (optimised model) a

	Initial	Extraction
SRec Correct	1.000	0.748
Visual Patterns test span	1.000	0.557
Spatial span	1.000	0.710
Spatial span (reverse)	1.000	0.637
Rey total A1-5	1.000	0.736
Rey A7	1.000	0.829
Rey Recognition A	1.000	0.745
Digit span (reverse)	1.000	0.665
FAS correct	1.000	0.476

The three components extracted explained 35.0%, 21.1% and 11.7% of the variance (35.0, 56.1, and 67.8% cumulatively).

Table 4-44. PCA #06 unrotated component matrix (optimised model) a

	Component		
	1	2	3
SRec Correct	0.580	0.142	0.625
Visual Patterns test span	0.669	0.318	-0.091
Spatial span	0.642	0.246	-0.487
Spatial span (reverse)	0.715	0.330	-0.128
Rey total A1-5	0.680	-0.521	-0.042
Rey A7	0.520	-0.738	-0.118
Rey Recognition A	0.495	-0.611	0.356
Digit span (reverse)	0.328	0.593	0.454
FAS correct	0.598	0.262	-0.223

a. 3 components extracted.

After Oblimin rotation, the pattern and structure matrices were again identical indicating an orthogonal solution which was confirmed with a subsequent varimax rotated model.

Table 4-45. PCA #06 Oblimin rotated pattern and structure matrices (optimised model) a

Pattern Matrix [a]

	Component		
	1	2	3
SRec Correct	0.044	-0.266	0.795
Visual Patterns test span	0.660	-0.030	0.203
Spatial span	0.872	-0.010	-0.203
Spatial span (reverse)	0.721	-0.035	0.185
Rey total A1-5	0.257	-0.778	-0.032
Rey A7	0.101	-0.877	-0.220
Rey Recognition A	-0.173	-0.837	0.262
Digit span (reverse)	0.189	0.281	0.718
FAS correct	0.674	-0.022	0.040

Rotation Method: Oblimin with Kaiser Normalization. a. Rotation converged in 7 iterations.

Structure Matrix

	Component		
	1	2	3
SRec Correct	0.305	-0.311	0.820
Visual Patterns test span	0.720	-0.154	0.382
Spatial span	0.819	-0.152	0.033
Spatial span (reverse)	0.777	-0.170	0.381
Rey total A1-5	0.383	-0.822	0.074
Rey A7	0.195	-0.884	-0.152
Rey Recognition A	0.043	-0.819	0.255
Digit span (reverse)	0.333	0.214	0.755
FAS correct	0.689	-0.141	0.223

Rotation Method: Oblimin with Kaiser Normalization.

Table 4-46. PCA #06 Varimax rotated component matrix (optimised model) a

	Component		
	1	2	3
SRec Correct	0.163	0.286	0.800
Visual Patterns test span	0.677	0.095	0.299
Spatial span	0.835	0.087	-0.072
Spatial span (reverse)	0.736	0.106	0.290
Rey total A1-5	0.311	0.799	0.019
Rey A7	0.143	0.879	-0.189
Rey Recognition A	-0.071	0.823	0.249
Digit span (reverse)	0.252	-0.247	0.735
FAS correct	0.670	0.086	0.140

Rotation Method: varimax with Kaiser Normalization. a. Rotation converged in 5 iterations.

As seen previously, this model is identical to the overall patient model with the SWM removed from the variable list.

4.4 Discussion of Chapter 4

The series of PCA above highlight the profile of neuropsychological tests/processes from the sample as a whole and individually for patients and controls. The approach to this analysis was to first examine the initial model which involved PCA with subsequent orthogonal (varimax) rotation. This was done in order to produce independent factors from which the composite scores can be derived. However, as discussed earlier, this approach may not be consistent with the real-life situation where it is entirely possible that these components of human neuropsychological functioning are indeed related. Therefore an optimised PCA was also performed using oblique rotation (oblimin) in order to examine variables on an item by item basis, excluding any that did not load uniquely or sufficiently onto an individual component. Once the pattern and structure matrices of these analyses converged, an orthogonal structure could be assumed and was confirmed by a return to a varimax method of rotation.

All analyses were also re-run from the initial data screening stage after excluding the variables from the CANTAB spatial working memory test. From the previous chapters, SWM is of particular interest given the results of Chapter 3 with the GR-antagonist mifepristone. One aim is therefore to examine the relationship between specific spatial tests (including SWM and the OLM test reported in the next chapter), composite neuropsychological factors and measures of HPA axis function.

4.4.1 Comparison of initial models (full variable set)

The initial varimax rotated models show a broadly consistent result between controls and the group as a whole. Two of the extracted components were identical – verbal learning and memory (component 2), and verbal executive function and WM (component 4). The remaining two components, containing different aspects of visuo-spatial memory processes,

were largely consistent although the order of extraction (and therefore proportion of variance explained) was reversed. The other difference was that in controls, the SRec variable was dropped from the model as it did not load significantly and consequently VPT which loaded across the two components reached the required level for inclusion in the strategic component only. SWM between search errors which loaded across both components in the overall sample only loaded onto one in controls, interestingly, with SOPT (another strategic self-ordered search task) and the SWM strategy score. Therefore this component in controls appears to be a 'cleaner', more distinct strategic processing factor. It is also worth noting that in the controls' PCA, the spatial span measures cluster with psychomotor speed in a separate component to the VPT. These tests are theoretically derived to tap different elements of the visuo-spatial sketchpad in Baddeley's WM model (Della Sala et al., 1999) – the inner-scribe and inner-eye respectively – and appear to cluster separately in healthy controls.

In contrast, a different profile of variable clusters emerged in the bipolar sample. Only three components were extracted and one of these (component 3) was somewhat weak, including only SRec and digit span (reverse)[21]. Moderate loadings were evident in this component, but did not reach the threshold for inclusion, from spatial span (reverse) and VPT. Therefore this method of orthogonal rotation may be less suitable for the patient sample due to a less independent structure underlying their neuropsychological processes. Component 1 contained a combination of visuo-spatial measures as well as verbal executive. The separation of the visuo-spatial sketchpad measures did not occur as it did in control participants and in component 2, the SOPT loaded along with the verbal learning and memory measures, possibly suggesting this task was being approached in a different way to controls, relying on verbal 'scaffolding' of performance. In fact, all three components included a combination of verbal

[21] It should be noted that reverse digit span has been shown to be impaired in visual neglect patients and although there is debate as to the underlying cause of this, both reduced general attentional processes and/or impaired spatial imagery have been suggested (Robertson, 1990; Rapport et al., 1994). Therefore the loading to a component with the spatial recognition task may represent an additional spatial component.

elements alongside visuo-spatial measures. Again, although not reaching the criteria for inclusion, SWM between errors did load moderately into the verbal learning and memory component (-0.416).

4.4.2 Comparison of optimised models (full variable set)

The next stage of the PCA analysis strategy was to look to optimise these models, comparing the pattern and structure matrices from an oblique rotation method and attempting to achieve equivalence between the two, a point at which an orthogonal structure can be assumed. The relative stability of the initial solutions was demonstrated by this approach as other than the exclusion of some variables that loaded across several components (sometimes not sufficiently to cross the cut-off for inclusion, but nevertheless showing moderate factor loading) the models remained largely unaltered with respect to the remaining variables.

In the total sample, three variables that appeared in the initial model (SRec, SOPT and the DSST) were excluded in the optimised model due to moderate loadings with multiple components. Consequently, the two visuo-spatial components collapsed together producing a simple three component solution for the optimised model, consisting of visuo-spatial processing (where CANTAB SWM measures loaded together with spatial span measures and the VPT), verbal learning and memory (containing verbal learning and delayed recall and recognition), and verbal executive function/WM (containing reverse digit span and verbal fluency measures). The optimised models for the separate groups also closely resembled the initial solutions: aside from the exclusion of VPT and reverse digit span in controls and the exclusion of SOPT in the bipolar sample, all resulting from shared variance with other components, the models remained the same. Although the changes are subtle, these consequent optimised models are of particular interest – in the bipolar sample, the SWM loads together with components of the visuo-spatial WM system that have been proposed to

be separable, namely the spatio/temporal and visual components (Della Sala et al., 1999), while in controls these WM components remained distinct. Therefore looking more closely at these visuo-spatial components and attempting a more detailed exploration of processes within is of interest (see chapter 5).

4.4.3 Comparison of models following SWM exclusion

In order to allow the compilation of composite scores that can be used as predictors of spatial working memory function, the SWM variables were removed from those available to the initial model and the PCA process was repeated. The initial model for the total sample produced a three component solution – two of these were identical to those produced in the models described above: a verbal learning and memory component and a verbal executive. The remaining factor was a general visuo-spatial component; this contrast with the two visuo-spatial components evident in the initial model with SWM included. Although variables such as DSST and SRec also dropped from the model, it is possible that the loading of SWM between search errors across both visuo-spatial components in the initial overall model (PCA #01) was the feature that kept these components separate and by removing it, the two collapsed together.

After separating by diagnostic group, the initial models with SWM removed again retained components identical to the initial PCA #03 and #05 models with respect to verbal memory and verbal executive although some changes occurred within the visuo-spatial components. In controls for example, with the SWM removed, the VPT incorporated with the component including spatial span and DSST, while the SOPT formed a new unique component with Pattern Recognition – a purer 'visual' component than had been seen. In patients, there were no changes to the components other than omitting SWM. Comparing the (SWM omitted) optimised with initial models, for the total sample and the patients, the only minor change to

occur was the removal of SOPT for loadings onto other components. In controls, the additional removal of the measures that formed the 'visual' component resulted in a reduction to a three component solution.

4.4.4 Summary

Overall, one of the most prominent features of the PCA models was the relative consistency between the components extracted, particularly the verbal learning and memory and verbal executive components. More variation was evident in some of the visuo-spatial components, with variables loading slightly differently when those exhibiting some degree of shared variance were removed. After separating by diagnostic group, the visuo-spatial components in controls remained divided into a more executive/strategic/cognitively-demanding element, and a shorter-term, immediate element. In patients, it was notable that overall only three components were ever produced and these showed a much greater overlap between verbal and visuo-spatial variables, with at least one verbal measure being evident in every component.

Very few studies have adopted a factor analytical or PCA approach to the assessment of neuropsychological processes in bipolar disorder. A study by Czobor and colleagues compared the neuropsychological factor structure of patients with bipolar disorder with patients with schizophrenia and reported a common six factors in both samples: attention, working memory, ideational fluency, verbal knowledge, non-verbal functions and learning (Czobor et al., 2007). However, within these factors there were some significant differences in the profiles of impairment between the diagnostic groups (patients with schizophrenia performing worse in the attention and non-verbal domains).

It is important to note the distinction between the PCA approach employed here and factor analysis. Factor analysis derives a mathematical model from which factors are estimated whereas PCA decomposes the available data into sets of linear variables. As such it has been argued that only factor analysis can truly estimate the underlying factors, with PCA simply examining the strength of the relationship between a given variable within each linear component, although the two approaches can lead to similar results when communalities are high (>0.7) (Field, 2000). As can be seen in the present analysis, a number of variables were excluded at the initial data screening stage and further removed from the model due to insufficient or multiple component loadings.

Finally, it is of particular interest to note the components that emerged from the optimised model in PCA#05 (the bipolar group). The first component, explaining the greatest proportion of variance in this model (38.2%), was a visuo-spatial memory/executive component. This included the variables: SWM BSE and strategy, spatial span forwards and reverse, VPT, and FAS verbal fluency. It is worth considering these in terms of the variables that were improved in the study in Chapter 3 following treatment with the GR antagonist mifepristone (SWM BSE, and improved from baseline, FAS and spatial recognition). Therefore, it is not simply the case that this general component and variables subsumed within were improved (although it is noted that not all tests from this component were used in Chapter 3). It may be that some process, separate from that which led to the clustering of variables within the PCA component, which is key to explaining which processes are changed by GR manipulation. As already mentioned above, it is also worth noting that in the bipolar sample, the SWM BSE loads together with components of the visuo-spatial WM system that have been proposed to be separable, namely the spatio/temporal and visual components (Della Sala et al., 1999).

Therefore looking more closely at visuo-spatial functioning and attempting a more detailed exploration of processes within this type of memory is important; firstly from the general point of view of gaining a better understand visuo-spatial memory processes in bipolar depression (see Chapter 5), and secondly, as a means of establishing the relationships between fractionated visuo-spatial memory processes, broader neuropsychological composites, and specific measures of HPA axis function related to the GR (see Chapter 6). Prior to this, the results of a novel memory paradigm which permits the fractionation of different spatial memory processes is reported. This task was administered to a sub-group of participants from the n=100 that took part in the present chapter.

Chapter V

Fractionation of spatial memory processes in bipolar depression

5. Fractionation of spatial memory processes in bipolar depression

5.1 Introduction to Chapter 5

Human spatial memory is far from being a unitary construct (for an overview see Schacter & Nadel, 1992). As already discussed in the case of the CANTAB SWM paradigm for example, there seems to be a clear distinction between the holding of spatial information 'online' over short periods of time and longer term maintenance (viz. within- and between-search errors). This can been seen both in behavioural processes and in underlying neural circuitry (see section 3.4.1). Within the working memory (WM) theory literature there is also increasing understanding of the independence of sub-processes within some WM slave systems, such as the fractionation of spatial/sequential components from a visual component within non-verbal short-term memory (Della Sala et al., 1997; Baddeley, 2000). One specific aspect of this latter component (i.e. memory for spatial layout/relationships between elements) has been further fractionated with reference to object-location memory where exact, metric (or 'coordinate') processes have been separated from relative relations between objects (or 'categorical' processes). A number of studies have now been carried out in healthy participants as well as patients with brain damage examining the separation of these processes using the same test paradigm – the Object Relocation test (ORT) program (Kessels et al., 1999)[22] – which can assess several discrete aspects of spatial memory and object binding, as well as their integration. As discussed in previous chapters, there are many hypothetical links between spatial memory, the HPA axis and mood disorders. However, to date, the assessment of these fractionated processes has not been explored in patients with bipolar disorder or depression.

[22] This paper is the first report of the paradigm as a complete computer program/software, in the form that was utilised in the present study. However, it should be noted that elements of the task were used in earlier papers as discussed subsequently.

The ORT program itself was developed as a standardised method of assessing different components of this specific form of spatial memory, namely the processes underpinning memory for spatial configurations and the localisation of objects within space (Kessels et al., 1999). The background to this task was developed through a series of studies by Postma and de Haan in healthy subjects, which sought to establish the independence of these processes. The initial study (Postma & de Haan, 1996) reported a series of three experiments in which participants were required to remember (from an initial 30 second exposure) and then relocate (as accurately as possible) 4, 7 or 10 objects into a 10.5cm x 10.5cm square frame on the screen. The objects were either verbal (i.e. letters), non-verbal (nonsense stimuli) or all identical, and these were recalled under two different task conditions – silently (AS-) or with articulatory suppression (AS+; counting backwards in 'ones' from 100). The main findings were that displacement error was lowest when all objects were identical ('position-only' condition) and just the spatial locations had to be remembered. Interestingly, performance on this condition was affected to only a limited extent by an increase in set size: after a slight decrease in performance from set-size 4 to 7, there was no subsequent difference when increasing to 10 and there was also no effect of AS+. Of the other two stimulus types, performance was very different – while accuracy was better overall for the placement of letters than nonsense objects, performance with both types was impaired by increasing set size and importantly, also by AS+ (i.e. the interference effect was not restricted to stimuli that were specifically verbal). From this data, there is the initial suggestion that a different process underlies memory for overall spatial layout compared to the relocation of objects within that space. In a second experiment, the spatial relocation phase was altered so that the exact positions in which objects had been located were given to the participant through the presentation of pre-marked locations. As before, the binding of letters to a location was easier than nonsense stimuli, and both set size and AS+ significantly impaired performance with both stimulus sets.

In one final experiment, the task was administered in a form very similar to that of the computerised version, with three conditions – the reconstruction of positions only (POM; position-only memory), the placement of objects to remarked locations (OLB; object-location binding), and a final condition that integrated both processes i.e. required participants to locate individual objects into a free space (COM; the combined condition). For this experiment, only set size 7 and 10 were used, the object stimuli used were a series of different punctuation marks (thereby falling somewhere between letters and nonsense stimuli in difficulty) and the articulatory suppression task was made less difficult in order to place fewer demands on overall cognitive resources (and now involved repeating the syllable 'blah'). Again, the relative independence of the POM process was demonstrated, with neither set size nor AS+ having any significant effect on performance. With both OLB and COM, both set size and AS+ significantly impaired relocation accuracy (although for COM, the AS+ only had an effect at set size 7).

Through this work, Postma and de Haan established preliminary evidence for a dissociation of (at least) two separable spatial processes in short term object location memory – the encoding of positions per se and object-to-position binding. Interestingly, the binding of objects to positions was shown to be sensitive to the effects of verbal articulatory suppression, even when non-verbal stimuli were used. One potential theoretical explanation for the separation of these processes was taken from the earlier work of Kosslyn and colleagues on spatial relationships used in visual perception and visuospatial imagery (Kosslyn et al., 1989; Kosslyn et al., 1992)[23]. Kosslyn had proposed that two types of processes underlying spatial representation exist – *co-ordinate*, which are involved in fine-grain, exact metric location, and *categorical*, which deal with more gross, relative relations between objects. It was also

[23] It should be noted that some authors have disputed the simple transfer of categorical/co-ordinate coding distinctions onto this form of short term visuo-spatial memory (Dent, 2009), specifically with regard to the fractionated processes. However, even here it is conceded that the debate is not over the independence of the processes but of the application of a categorical/co-ordinate distinction to them.

proposed that there was *relative* hemispheric specialisation for these processes, with categorical judgments being faster when stimuli are initially presented to the left cerebral hemisphere, compared to evaluations of distance, which are faster when initially presented to the right hemisphere. There was further suggestion that these categorical representations developed with practice and therefore the possibility of verbal coding of such relations was suggested, although it is noted throughout this work that it is unlikely to be verbal/linguistic labelling *per se* but a commonality of process behind coding of information in a categorical or modular way (Kosslyn et al., 1989). In work conducted earlier to Postma and de Haan (1996) it had been proposed that this distinction between categorical and co-ordinate processes may extend from early perceptual systems and apply also to long-term memory (McNamara, 1991). The work by Postma and de Haan had extended this to processes within short-term visuo-spatial memory which led to a series of subsequent studies exploring this phenomena in healthy participants and those with brain lesions or diffuse impairment, as well as examining some of the neurobiological modulators of these processes (e.g. the role of sex hormones). A brief overview of this work is now presented.

5.1.1 Effects of sex hormones on the ORT paradigm

An extensive literature exists exploring sex differences in spatial abilities and the underlying neurobiology (Geary, 1995; Halpern & Wright, 1996; Cahill, 2006). The ORT has been used in a number of studies to examine overall differences in fractionated processes as well as the effect of AS+ on performance (Postma et al., 1998; Postma et al., 1999). In both these studies, males were found to be more accurate at POM, with no sex difference in OLB or COM performance (although in the earlier study, using a 'best-fit' error score rather than absolute error yielded a male superiority). In both studies, contrary to expectations, AS+ reduced performance in all conditions (which the authors suggested may be the result of a combination of factors such as stimulus set size and overall frame size - 15 cm x 15 cm was

used). In the 1999 study, the role of hormonal factors was also examined and it was found that POM performance was significantly worse in females at the point of the menstrual cycle when oestrogen levels are lower (Postma et al., 1999). In a follow-up study, the effect of testosterone administration (sublingually 0.5 mg testosterone or placebo with cyclodextrin as carrier) was examined in female healthy volunteers (Postma et al., 2000). This study included both an immediate and a delayed recall version of the task (using a 3 min retention interval). It was found that for delayed recall of the COM condition, performance was significantly improved by testosterone.

Although tentative, the effects of these hormones at the level of the hippocampus was suggested as a potential mechanism through which these effects emerged. This has implications for the use of the task in the present study as it has been shown that some of the effects of sex hormones on memory function may be via interactions with corticosteroids (Symonds et al., 2004) and there is increasing understanding of these interactions of hippocampal morphology and function (for a review see McEwen, 2010).

5.1.2 Effects of brain injury on the ORT paradigm

A further important source of information on the brain structural underpinnings of spatial memory processes is from a series of studies examining performance on the ORT paradigm in patients with brain damage. One of the first studies (Kessels et al., 2000a) assessed 10 patients after intracranial tumour resection (compared to 24 healthy controls). Five patients showed no impairment and one was impaired on all experimental levels of the task. Interestingly of the remaining patients a double dissociation was observed, with two patients being impaired on POM only and two on OLB and COM, but not POM. These latter two patients who were impaired on OLB and COM were also significantly impaired on the object memory control task. Although it was suggested that these general memory deficits cannot

fully explain the selective spatial memory findings, as it is especially spatial mnemonic processing that is tested, rather than memory for object identity (i.e. patients do not have to recall the objects), it has been shown in several studies that damage to the parietal lobes can cause impairments in visual discrimination and visual-manual exploration which may have affected performance (Eacott & Gaffan, 1991; Hinkley et al., 2009).

Due to the varying extent and localisation of the surgery in the patients in the above study, it was not possible to be precise about the underlying neural circuitry of the spatial processes examined although overall spatial memory impairment was more frequent in right hemispheric patients than in left hemispheric patients and was more often found in patients with posterior (parietal or occipital) lesions than in patients with anterior (temporal or frontal) lesions (Kessels et al., 2001). In a larger follow up study (Kessels et al., 2002) in patients who had suffered a stroke with left hemisphere infarct (LH; n=28), right hemisphere infarct (RH; n=16) or bilateral (BIL; n=6) similar left-right distinctions were observed with RH patients being (statistically) significantly impaired at POM only (immediate and delayed recall) compared to controls, while LH patients only were impaired at OLB (immediate and delayed recall) and COM (immediate recall).

In subsequent studies, it has been possible to characterise and group patients more precisely in terms of lesion location. Kessels et al (2004) administered the ORT to twenty five patients who had suffered from medically refractory temporal-lobe epilepsy caused by mesiotemporal sclerosis and had undergone a unilateral selective amgygdalohippocampectomy (16 left side AH, 9 right side AH). The task included an additional condition – a categorical POM trial – in which a grid was included in the relocation phase of identical stimuli. Overall, right AH patients were selectively impaired on POM while left AH patients were impaired on COM, compared to controls. There was no significant difference on the categorical POM trial,

although for the OLB condition, a trend was observed in the group effect, with left AH patients making more errors.

In two final studies (van Asselen et al., 2008; van Asselen et al., 2009), recently conducted in stroke patients with LH (n=13) or RH (n=12) damage, van Asselen and colleagues tested more specific hypotheses using the ORT paradigm, focussing more on the categorical/co-ordinate distinction in memory for exact spatial locations versus the binding of objects to those locations. Effects were observed on positional memory only, with LH patients being impaired on the categorical process while RH patients were impaired on coordinate processes (van Asselen et al., 2008). Using a lesion-localisation method in a larger group of patients they then went on to demonstrate that the area of maximal lesion overlap for the POM task condition was in the right hemisphere (including the insula, the superior/middle temporal cortex, the posterior parietal cortex and the inferior frontal gyrus); for OLB the area of maximum overlap involved the left posterior parietal cortex and the right hippocampus, putamen and fusiform/lingual gyrus; and for COM the overlap also involved the left posterior parietal cortex (van Asselen et al., 2009).

5.1.3 Implications for the subsequent study

Although there are some minor differences in which processes are affected by lateralised brain damage (most likely due to differences in the cause and type of damage), the overall profile is relatively consistent. As discussed previously, some authors have questioned the direct application of categorical and coordinate processing explanations to the binding of objects to spatial locations, although this disagreement relates only to the mechanism and not to the dissociation of memory for positions from memory for object-location binding (Dent,

2009)[24]. In general it appears that there are three components to this form of visuo-spatial memory – memory for objects, spatial locations, and the binding of objects to those locations (Postma et al., 2008). Focussing on the latter two spatial processes, it appears that memory for exact metric locations is dependent upon the RH, especially posterior parietal cortex. The binding of objects to locations seems to be dependent on the LH with an important role for the hippocampus bilaterally (Piekema et al., 2006; Postma et al., 2008; van Asselen et al., 2009)[25]. Along with the evidence from healthy volunteer studies that some of these processes can be modulated by changes in some exogenous and endogenous steroids, the ability to dissociate these components using the ORT paradigm provides an important method to explore visuo-spatial memory in more detail in the context of the present research. To date, the paradigm has not been used in patients with depression or bipolar disorder (although in schizophrenia, the COM process has been examined in isolation and found to be impaired, but only when stimuli with threatening content were used van 't Wout et al., 2007).

The aim of this chapter is therefore to examine the performance of patients with bipolar depression and healthy controls on a this novel visuo-spatial memory task – the ORT paradigm. The participants were part of the initial cohort (sequentially recruited) from those taking part in Chapter 4 therefore all additional secondary neuropsychological measures are available to compare performance on the fractionated ORT processes with.

[24] However, there are many differences between the paradigm used by Dent (2009) and the ORT paradigm which leaves this debate open, such as the use of only 4 stimuli-location pairings per trial, the use of colours rather than nameable objects and the use of a rapid change detection methodology. Other work has shown a clear time-course effect whereby LH categorical advantage increases over retention interval suggesting that the maintenance of representation in this way may be linked to efficiency of coding over fine-grain detail (Postma et al., 2006).

[25] However it is important to note that aspects of these distinctions may be relative, for example in their review Postma and colleagues note that there are frontal and hippocampal contributions to memory for exact metric locations which are time-dependant i.e. greater hippocampal involvement when longer term maintenance is required (Postma et al., 2008).

5.2 Methods

As stated previously, these participants were a sub-set of those included in chapter 4 (n=100). The following methods have been previously outlined and are therefore presented only in brief.

5.2.1 Subjects

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. 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 standardized 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.

5.2.2 Neuropsychological tests

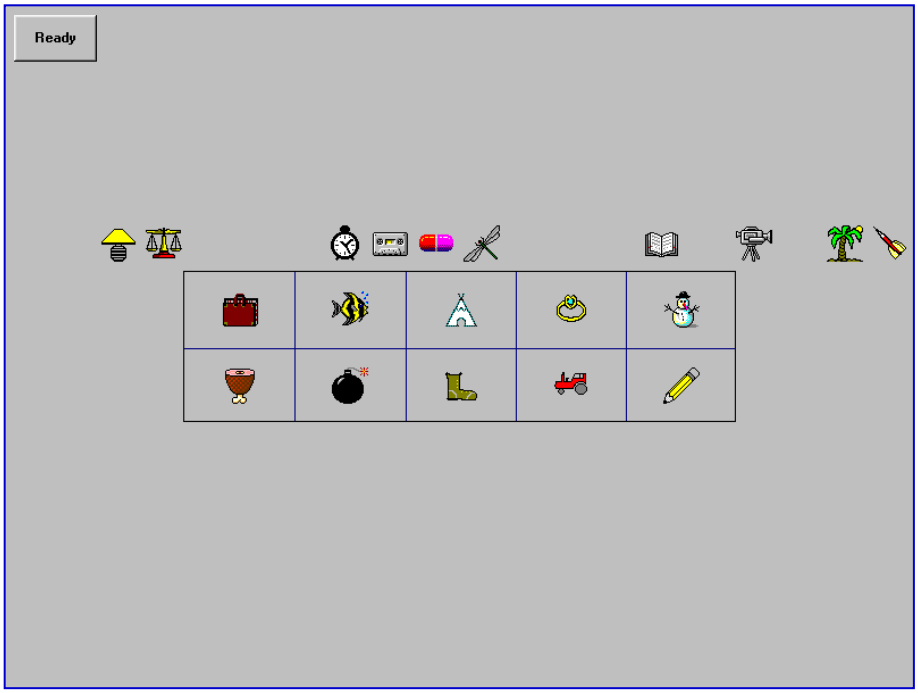
5.2.2.1 Object Location Memory

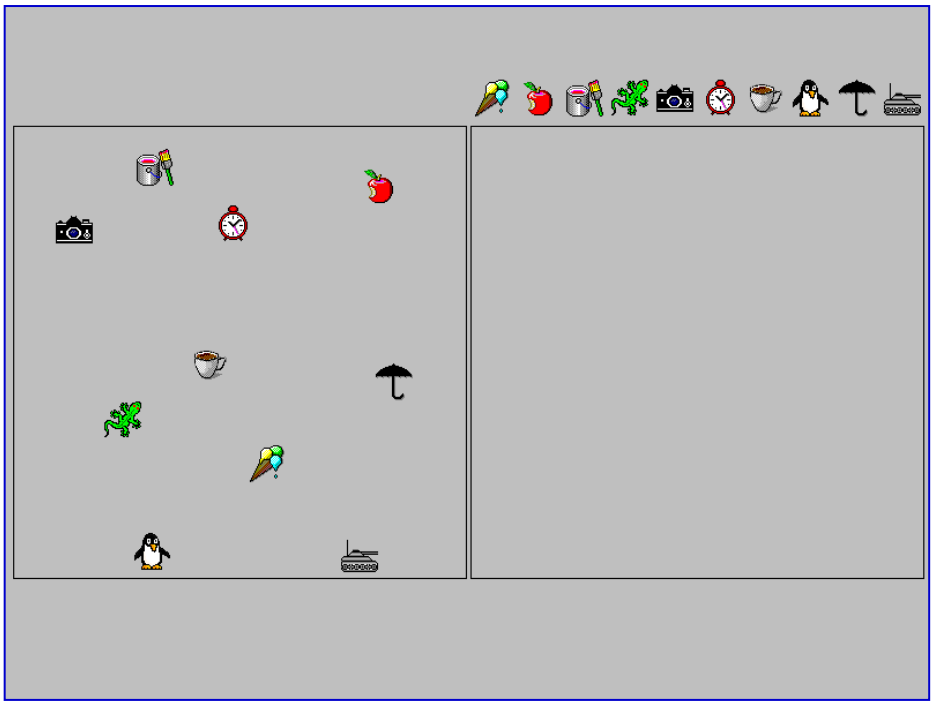
To assess memory for the locations of objects, the Object Relocation program was used (Kessels et al., 1999). The program presents stimulus displays on a PC fitted with a touch-

screen monitor. A number of variations of the task parameters are possible within the program (e.g.; Kessels et al., 2004); here we ran the program using the immediate memory conditions from (Kessels et al., 2000a).

First, subjects completed two control tasks that assessed *object identity memory*; and *visuospatial construction and perception*. In the object identity task subjects viewed 10 different objects for 30s which had to be remembered and subsequently recognised from a set of 20 objects, containing the ones that were shown previously and 10 distracters. In the visuospatial construction task subjects had to copy a frame containing 10 different objects at different locations without a memory component. Each task condition consisted of an example containing only 4 objects/positions, followed by two different test displays.

Figure 5-1. Control conditions for the OLM task (upper figure shows the object identity trial and lower the visuospatial reconstruction)





Following these control tasks, subjects completed 3 experimental task conditions:

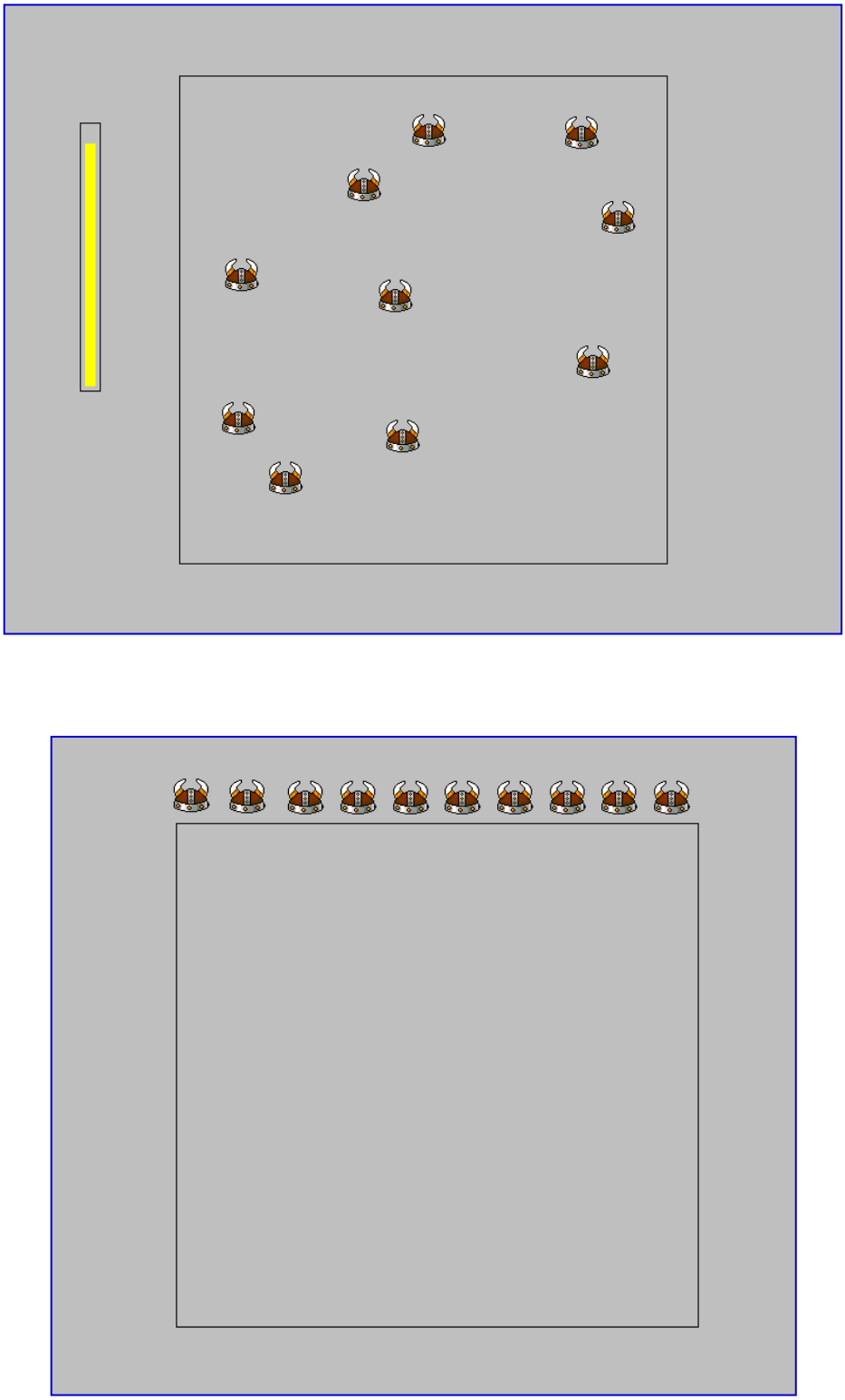
(i) *position-only memory (POM)* - subjects viewed an array containing 10 identical objects and were required to remember their precise locations. After 30s the array disappeared and the objects appeared along the top of the screen. Subjects were then required to move the objects down into the empty frame and recreate the exact positions of the array as accurately as possible.

(ii) *object-location binding (OLB)* – subjects viewed an array of 10 different objects and were required to remember where they were located within the frame. After 30s the array disappeared and the objects appeared along the top of the screen. Subjects were then required to move the objects down into the frame and recreate the array, although the precise positions that had been occupied were indicated by pre-marked by black dots.

(iii) *combined memory condition (COM)* – which was identical to the OLB condition except for the relocation stage where there were no pre-marked black dots i.e. subjects were required to remember and relocate the 10 different objects as precisely as possible to their exact previous locations.

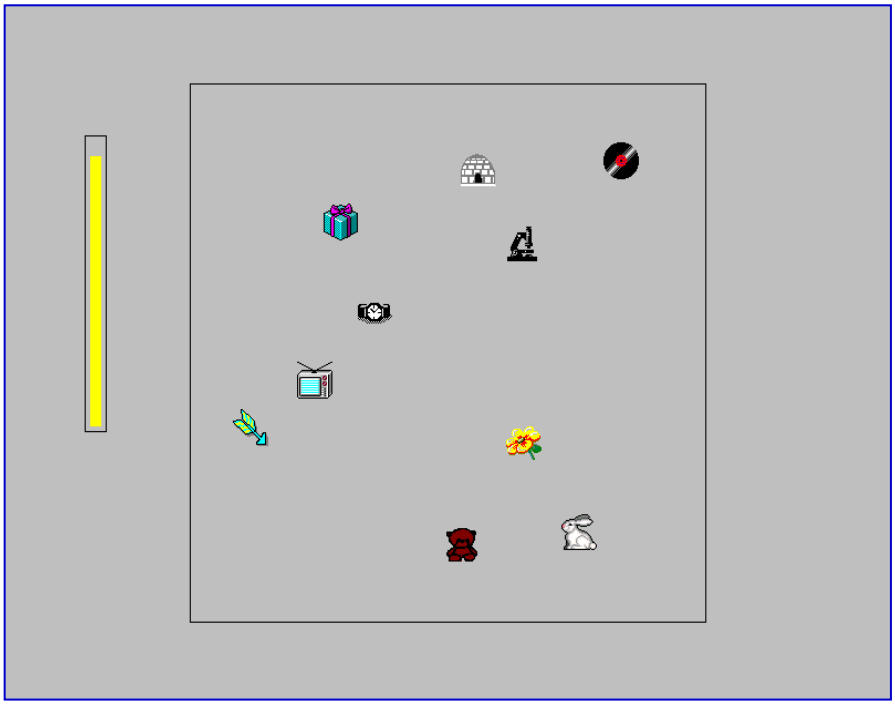
Again, each task condition consisted of an example containing only 4 objects/positions, followed by two different test displays. Performance measures were percentage incorrect items in the object identity control condition and OLB conditions, and deviation error (millimetres; mm) in the visuospatial construction and perception control condition, and POM and COM tasks. In the case of the POM task, as all objects are identical, it is impossible to specify which location any given object is relocated to and consequently the best-fit error is used (Kessels et al., 2000a). All other tasks use the absolute error score.

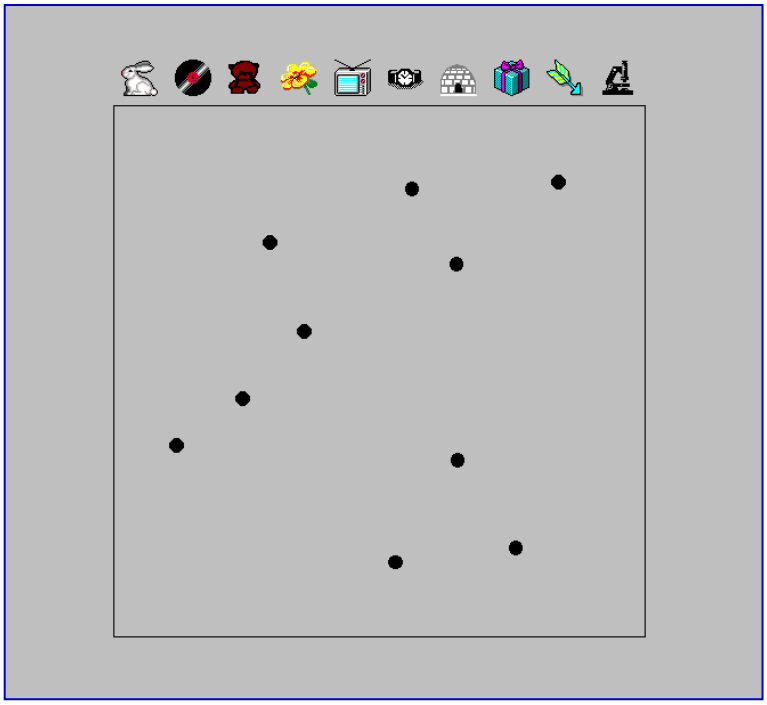
Figure 5-2. POM; position-only memory condition (upper figure shows the learning phase and the lower figure the test phase)



Nb. Figures not to scale

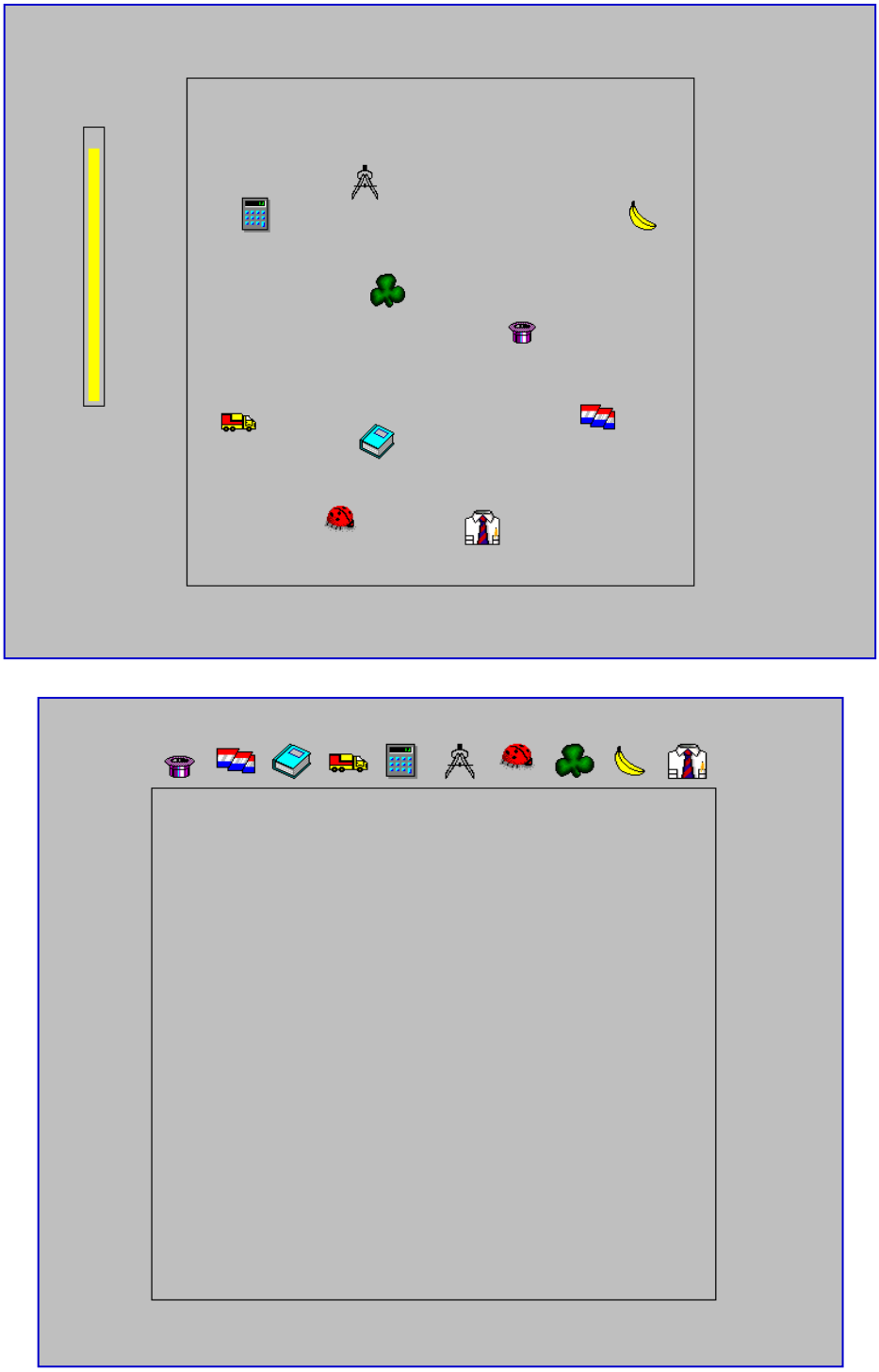
Figure 5-3. OLB; object location binding condition (upper figure shows the learning phase and the lower figure the test phase)





Nb. Figures not to scale

Figure 5-4. COM; combined condition (upper figure shows the learning phase and the lower figure the test phase)



Nb. Figures not to scale

5.2.2.2 Secondary neuropsychological tests

Additional secondary neuropsychological tests (with a main focus on visual and spatial processes) were also administered. These have been outlined in previous chapters (chapter 4) and are only briefly listed below.

CANTAB Spatial Working Memory (SWM): a self-ordered search task which requires subjects to search for hidden tokens within a spatial array. The number of between-search errors are recorded i.e. occasions when a subject returns to a square under which a token was already found, as well as a strategy measure, were a lower strategy score reflects a more systematic search strategy. *CANTAB Spatial Recognition (SRec)*: a memory task in which subjects view 5 'squares' presented in serial order and then are subsequently required to identify, from a choice of 2 squares, the one that occupies one of the 5 locations shown previously. Subjects complete 4 sets. The percentage of correct responses are recorded. *CANTAB Spatial Recognition-modified (SRec-m)*: a modified version of the task was also administered which is identical to the standard version except two sets of 7 squares, then 2 sets of 9 squares are used. *CANTAB Spatial Span and Reverse Spatial Span (SSp/ rSSp)*: a test analogous to the Corsi Block task which is administered first in the standard format and then reverse, where subjects tap the sequence in the opposite order from presentation. The maximum span reached is recorded. *Visual Patterns Test (VPT)*: a test of short-term visual memory in which subjects are required to remember and reproduce increasingly complex 'checkerboard' patterns (Della Sala et al., 1999). It is scored in the same way as the SSp task with the maximum set-size achieved being recorded. *CANTAB Pattern Recognition (PRec)*: a test of visual recognition memory in which subjects view a series of 12 coloured patterns and must then select the patterns they have seen in a 2-choice, forced-discrimination paradigm. Subjects complete 2 sets and the percentage correct is recorded. *Pattern Recognition-modified (PRec-m)*: due to the risk of ceiling effects in healthy controls, a modified pattern recognition task was constructed which

was similar to the CANTAB version except the patterns were more abstract, black and white shapes and were more closely matched to their distracter during the recognition phase. These were taken from (Vanderplas & Garvin, 1959) and displayed using the Superlab program. One set of 24 patterns was administered. *Self-Ordered Pointing Test (SOPT)*: a test of executive function/ visual working memory, requiring the ability to generate and monitor a sequence of responses. The version used here consists of 3 trials at levels 4, 6, 8 and 10.

Rey-Auditory Verbal Learning Test (Rey-AVLT): a verbal learning and memory task which was administered according to standardised instructions (Rey, 1964; Lezak et al., 2004). *Forward and Backward Digit Span (fDSp/ bDSp)*: a test of immediate verbal recall and working memory which was again administered according to standardised instructions (Lezak et al., 2004). *Verbal fluency (FAS and ELFT)*: tests of verbal fluency with difference executive demands. *Digit Symbol Substitution Test (DSST)*: testing psychomotor speed. Lastly, speed of cognitive processing (*SCOLP; Speed and Capacity of Language Processing*) was included.

5.2.3 Statistical analysis

Descriptive statistics are presented as mean, standard deviation (s.d.) and range. For graphical presentation of results, bar charts are presented as mean, with error bars representing $\pm$ 1 standard error of the mean (SEM). Where estimates effect sizes are presented Cohen's d is used (Cohen, 1988), calculated using the formula ($\overline{X}_{\text{patients}}$ minus $\overline{X}_{\text{controls}}$)/$S_{\text{pooled}}$). Where necessary the signs of the effect sizes were reversed so that negative values always represent impairment in the patient group compared to the controls. For the primary analyses, a parametric approach was again adopted. The secondary battery was subjected to an overall multivariate analysis first for consistency (see section 2.2.6) however the primary aims of this chapter was to examine the performance of participants on the ORT and examine the relationship to the secondary measures. Correlation and hierarchical regression (entry

method) were adopted. Differential deficit in ORT measures was examined using the program developed by Crawford and colleagues ('diffdef'), which can be used to test whether the deficit exhibited by a clinical sample on one measure is significantly greater than the deficit exhibited on another through the application of William's (1959) test for non-independent correlations (Crawford et al., 2000). Statistical analyses were carried out using SPSS version 17 (SPSS, 2008).

5.3 Results

Twenty-five patients (n=17 male) and 25 healthy controls (n=19 male) took part in the study (these were sequentially the first participants from the n=100 from chapter 4). The two groups were well matched by sex (chi-squared=0.97, *df*=1, *p*=0.529), age (BD: mean=46.1 years, s.d.=10.9; controls: mean=44.2 years, s.d.=15.1; *t*=0.515, *df*=48, *p*=0.609), years of education (BD: mean=13.9 years, s.d.=2.5; controls: mean=14.5 years, s.d.=2.3; *t*= -0.829, *df*=48, *p*=0.411) and NART estimated full-scale IQ (BD: mean=110.9, s.d.=9.9; controls: mean=112.0, s.d.=13.2; *t*= -0.329, *df*=48, *p*=0.744).

Patients had a mean age of illness onset of 30.2 years (s.d.=13.1) and a current median length of illness episode of 18 weeks (mean=51, s.d.=74). Severity of depression in the group at screening using the HAM-D17 was 19 (s.d.=4.4) and on the day of testing (using MADRS) was 26 (s.d.=8.5). The median number of hospitalizations in the group was 3.

5.3.1 Object Location Memory

On the experimental conditions, patients with BD performed significantly worse than controls on all 3 measures: OLB (t=2.611, df=48, p=0.012; d= -0.70), POM (t=4.169, df=48, p<0.0001; d= -1.02) and COM (t=2.987, df=48, p=0.004; d= -0.78). Application of Crawford's calculator did not reveal evidence of a differential deficit between any of these processes (POM vs. OLB: t=1.181, df=47, p=0.243; POM vs. COM: t=0.846, df=47, p=0.402; OLB vs. COM: t=0.380, df=47, p=0.706).

Examination of the control conditions revealed that while performance of object identity memory did not differ significantly between groups (t=1.063, df=48, p=0.293; d= -0.30), patients with BD performed significantly worse than controls at the visuo-spatial construction task (t=3.120, df=48, p=0.003; d= -0.81).

Table 5-1. Mean (s.d.) errors for patients and controls on the Object Location Memory test

	Patients (n=25)		**Controls (n=25)**	
	Mean	s.d.	Mean	s.d.
Control tasks				
Object identity (% errors)	4.4	(9.3)	2.2	(4.6)
Visuospatial construction (error, mm)	107.2	(51.2)	69.8	(31.2)
Experimental measures				
OLB (errors, %)	34.0	(22.9)	17.8	(21.0)
POM (error [a], mm)	200.5	(49.7)	150.5	(33.5)
COM (error, mm)	310.7	(83.1)	239.2	(86.3)

[a] *in the case of POM, the best-fit error was used (see methods section).*

An ANCOVA was therefore used to examine the group differences on each of the experimental measures, with the inclusion of visuospatial construction as a covariate (this method has been utilised in previous studies, see van Asselen et al., 2008). Sex was also added as a between subjects factor. Patients were significantly impaired on the POM measure (F=6.733, df=1,45, p=0.013). However performance on the OLB measure (F=1.918, df=1,45, p=0.173) and COM measure (F=3.635, df=1,45, p=0.063) did not significantly differ between groups. The visuo-spatial reconstruction task was a significant covariate in all 3 models. On the OLB measure, a significant sex difference was observed with females performing significantly better than males (F=5.351, df=1,45, p=0.025).

Although there was no differential deficit, examination of the relationship between the three measures revealed that performance on the OLB and COM measures was significantly correlated for patients (r_s=0.521, p=0.008) and controls (r_s=0.550, p=0.004), but there was no significant correlation between POM and either of these measures in patients (POM vs. OLB: r_s=0.334, p=0.103; POM vs. COM: r_s=0.251, p=0.227) or controls (POM vs. OLB: r_s=0.085, p=0.688; POM vs. COM: r_s=0.192, p=0.359). This suggests that POM may tap different underlying processes than OLB and COM.

5.3.2 Secondary neuropsychological tests

As discussed in chapter 4.1.1, ceiling effects were observed in the standard CANTAB PRec test therefore only the modified PRec is included in the multivariate analysis. For the SRec, the standard version was used (5 stimuli). Also, as the groups were intentionally matched on pre-morbid estimated verbal IQ, the spot the word test from the SCOLP was not included in between group comparisons.

From the primary MANOVA, a overall main effect of group was observed (Pillai's Trace=0.573; F=1.945, df=20,29, p=0.050) with patients performing worse than controls. Examination of individual tests is presented in Table 5-2 indicating statistical differences across the majority of measures included.

Table 5-2. T-test for between group differences for patients and controls

	Patient		Control		t-test		ES [a]
	mean	s.d.	mean	s.d.	t	p	d
Visual Patterns test							
span	7.9	(1.6)	9.3	(2.2)	-2.584	**0.013**	-0.69
SOPT							
total errors	13.5	(5.5)	10.0	(6.3)	2.096	**0.041**	-0.57
Pattern Recognition							
correct (standard)[b]	21.3	(2.8)	22.3	(2.2)	-1.409	0.168	-0.40
correct (modified 24)	17.1	(3.0)	18.9	(2.5)	-2.285	**0.027**	-0.62
Spatial span							
forward span	5.2	(1.0)	6.0	(1.2)	-2.470	**0.017**	-0.66
reverse span	5.2	(1.1)	6.2	(1.4)	-2.865	**0.006**	-0.76
Spatial Working Memory							
between errors	31.6	(20.1)	22.7	(18.8)	1.701	0.092	-0.45
within errors	2.7	(7.5)	1.3	(1.8)	0.897	0.374	-0.25
strategy score	33.4	(6.4)	30.7	(6.2)	1.552	0.127	-0.43
Spatial Recognition							
correct (standard)	13.7	(3.0)	14.6	(3.1)	-1.060	0.294	-0.30
correct (modified 7) [b]	9.1	(2.2)	10.4	(2.2)	-2.010	**0.050**	-0.55
correct (modified 9) [b]	11.1	(2.3)	11.3	(2.0)	-0.390	0.698	-0.11
Rey-AVLT							
correct (total A1 to A5)	38.7	(8.6)	46.0	(8.6)	-3.025	**0.004**	-0.79
correct (A7)	6.0	(3.4)	8.4	(3.5)	-2.459	**0.018**	-0.66
correct (recognition A)	11.5	(2.6)	12.1	(2.6)	-0.732	0.468	-0.21
Digit span							
forward span	6.2	(1.1)	7.3	(1.1)	-3.644	**0.001**	-0.92
reverse span	4.7	(1.2)	5.2	(1.3)	-1.465	0.149	-0.41
Verbal fluency							
'FAS' correct	37.8	(9.2)	44.4	(10.9)	-2.299	**0.026**	-0.62
'exclude letter' correct	34.9	(8.6)	45.1	(10.7)	-3.726	**0.001**	-0.94
DSST							
Correct	46.6	(10.2)	56.0	(9.8)	-3.330	**0.002**	-0.86
SCOLP							
Speed of Comprehension	55.2	(15.3)	71.9	(15.8)	-3.810	**<0.001**	-0.95

[a] *Effect sizes (ES) are Cohen's d, corrected so that negative values always represent impairment in patients compared to controls.*

[b] *these are included for information only.*

5.3.3 Relationship between the ORT and secondary neuropsychological tests

Presented below are the (Spearman's) correlation matrices for controls and patients separately, for the ORT task and the tests from the wider secondary battery (statistically significant correlations are highlighted in red; $p<0.05$).

Table 5-3. Correlation matrix for ORT measure with the secondary battery (Controls)

Control	POM	OLB	COM	VPT	SOPT	PRec m	SSp	SSpr	SWM ber	SWM wer	SWM strat	SRec	Rey total	Rey A7	Rey Rec A	Dspa n-f	Dspa n-r	FAS	ELFT	DSST
(a) VPT	-.681	.099	-.036																	
(b) SOPT	.280	.135	.117	-.433																
(c) PRec m	-.588	-.481	-.659	.406	-.183															
(d) SSp	-.470	.148	-.105	.505	-.481	.261														
(e) SSpr	-.377	.226	-.061	.265	-.188	.039	.422													
(f) SWM ber	.463	.163	.338	-.507	.552	-.407	-.654	-.389												
(g) SWM wer	.278	.062	.478	-.346	.289	-.441	-.445	-.370	.572											
(h) SWM strat	.187	.282	.436	-.477	.545	-.529	-.238	-.014	.620	.442										
(i) Srec	-.429	-.335	-.448	.506	-.421	.664	.555	.314	-.635	-.526	-.667									
(j) Rey total	.040	.039	-.365	-.078	-.312	.127	.166	.293	-.363	-.041	-.301	.238								
(k) Rey A7	-.104	-.054	-.360	.099	-.326	.180	.275	.275	-.480	-.284	-.302	.384	.799							
(l) Rey Rec A	-.123	-.148	-.522	.197	-.419	.351	.428	.232	-.593	-.606	-.395	.446	.524	.649						
(m) Dspan-f	-.044	-.113	-.041	-.082	-.122	-.194	.006	.335	-.175	.283	-.015	.034	.396	.229	-.023					
(n) Dspan-r	-.116	-.370	-.247	.098	-.605	.232	.190	.207	-.370	-.203	-.482	.392	.215	.263	.306	.314				
(o) FAS	.250	-.445	-.272	-.174	-.103	.105	-.236	-.249	.268	.215	-.075	.169	.132	.122	.019	.029	.329			
(p) ELFT	.107	-.535	-.079	-.056	.015	.188	-.370	-.060	.149	.212	-.180	.149	.007	.103	-.063	.098	.388	.544		
(q) DSST	-.180	-.309	-.370	.087	.016	.291	.203	.353	-.492	-.353	-.239	.321	.332	.421	.486	.181	.093	-.081	.300	
(r) SpofCom	.240	-.209	-.096	-.006	.160	.055	-.157	.038	.086	.132	-.155	.165	-.072	-.227	.023	-.064	.052	.513	.508	.194

Table 5-2. Correlation matrix for ORT measure with the secondary battery (Patients)

Bipolar	POM	OLB	COM	VPT	SOPT	PRec m	SSp	SSpr	SWM ber	SWM wer	SWM strat	SRec	Rey total	Rey A7	Rey Rec A	Dspa n-f	Dspa n-r	FAS	ELFT	DSST
(a) VPT	-.334	.059	-.070																	
(b) SOPT	.420	.177	.360	-.568																
(c) PRec m	-.527	-.520	-.519	.335	-.486															
(d) SSp	-.305	-.327	-.361	.428	-.604	.393														
(e) SSpr	-.540	.036	-.088	.638	-.633	.176	.645													
(f) SWM ber	.136	.233	.336	-.350	.416	-.257	-.650	-.529												
(g) SWM wer	-.288	-.13	-.152	-.101	.032	.222	-.157	-.106	.274											
(h) SWM strat	-.126	.125	.182	-.156	.262	-.160	-.414	-.216	.761	.178										
(i) Srec	-.243	-.026	.069	.084	-.221	-.006	-.027	.032	.053	.250	-.035									
(j) Rey total	-.060	-.523	-.404	.107	-.362	.219	.189	.070	-.176	.040	-.099	.067								
(k) Rey A7	.010	-.246	-.086	.271	-.327	.112	.278	.311	-.410	-.276	-.359	-.012	.527							
(l) Rey Rec A	-.241	-.370	-.248	-.053	-.243	.253	.021	.016	-.206	.130	-.236	.357	.225	.583						
(m) Dspan-f	-.132	-.176	-.229	.176	-.210	.007	.463	.460	-.250	.145	-.180	-.276	.011	-.102	-.286					
(n) Dspan-r	.278	.003	-.023	.126	-.059	-.242	.086	-.108	-.048	-.133	-.120	.207	.003	-.190	-.054	.033				
(o) FAS	-.071	-.073	-.430	.467	-.520	.310	.566	.315	-.265	-.022	-.357	.026	.096	.017	-.055	.411	.438			
(p) ELFT	.022	.106	.046	.336	-.262	.019	.283	.196	-.327	-.035	-.578	.231	.264	.403	.139	.023	.071	.343		
(q) DSST	-.278	-.348	-.275	.252	-.284	.345	.099	.144	-.180	.080	-.121	.149	.474	-.014	-.073	.184	.142	.177	.267	
(r) SpofCom	-.109	-.014	.030	.336	-.139	.035	.287	.159	-.160	.044	-.108	.213	.123	-.082	-.235	.240	.173	.310	.456	.592

Overall it can been seen that there appears to be a greater extent of inter-correlation between secondary tests in control subjects compared to patients, especially within the visuo-spatial tasks (rows a to i). When examining the relationship between the ORT paradigm outcome measures and the secondary measures, it is clear that a very different profile characterises each group. Aside from the PRec-m measure, there are no shared areas of significance of correlations between patient and control groups. With regard to the ORT measures, there are two areas are of particular note. Firstly, in controls, there are multiple significant correlations between the POM and the visuo-spatial tasks, including the spatial span and VPT measures, PRec-m and SRec, and SWM (between errors). There are numerically fewer in patients (although of note is the correlation between COM and SWM between errors). Secondly, in patients, there are significant correlations between verbal learning (Rey-AVLT total) and the ORT measures of OLB and COM, but not POM. Comparing the strength of these ORT/Rey-AVLT correlations between patients and controls revealed that there was no significant difference for POM (z=0.33, p=0.741) or COM (z=-0.15, p=0.881) measures, but the OLB correlation was significantly larger in patients than controls (z=-2.05 p=0.040).

Using a hierarchical multiple regression method, each of the ORT variables was examined separately, entering visuospatial reconstruction (VSR), Rey-AVLT total, and group as independent predictors. Four models were examined; Models 1 and 2 examined the individual effects of Rey-AVLT total and visuospatial reconstruction respectively by entering each of these variables first followed by group. Models 3 and 4 examined the order of entry.

Table 5-4. POM hierarchical regression (whole group, with Rey-AVLT total and VSR)

	R^2	R^2 change	F for R^2 change	p
Model 1				
Rey-AVLT total	0.090	0.090	4.755	**0.034**
Group: BD vs. control	0.276	0.186	12.094	**0.001**
Model 2				
VS reconstruction	0.227	0.227	14.056	**<0.001**
Group: BD vs. control	0.350	0.123	8.913	**0.004**
Model 3				
Rey-AVLT total	0.090	0.090	4.755	**0.034**
VS reconstruction	0.253	0.163	10.261	**0.002**
Group: BD vs. control	0.353	0.100	7.082	**0.011**
Model 4				
VS reconstruction	0.227	0.227	14.056	**<0.001**
Rey-AVLT total	0.253	0.027	1.679	0.210
Group: BD vs. control	0.353	0.100	7.082	**0.011**

From the regression analysis of the POM measure, entry of either variable (Rey-AVLT total or VSR) individually predicted a significant proportion of the variance, although the entry of group membership also explained significant additional variance (18.6% and 12.3% respectively). Importantly, when both variables were entered, the order of entry had a significant impact on the resulting models and although both explained a significant proportion of the variance (model 3), when VSR was added first (25.3%), the addition of the Rey-AVLT did not result in a significant increase (2.7%). Again, entering the group variable was significant.

A different pattern emerged for both the OLB and COM measures, with a greater proportion of the variance being explained by the verbal learning measure.

Table 5-5. OLB and COM hierarchical regression (whole group, with Rey-AVLT total and VSR)

OLB	R^2	R^2 change	F for R^2 change	p
Model 1				
Rey-AVLT total	0.186	0.186	10.953	**0.002**
Group: BD vs. control	0.224	0.039	2.341	0.133
Model 2				
VS reconstruction	0.119	0.119	6.478	**0.014**
Group: BD vs. control	0.172	0.054	3.042	0.088
Model 3				
Rey-AVLT total	0.186	0.186	10.953	**0.002**
VS reconstruction	0.236	0.050	3.101	0.085
Group: BD vs. control	0.253	0.017	1.034	0.315
Model 4				
VS reconstruction	0.119	0.119	6.478	**0.014**
Rey-AVLT total	0.236	0.117	7.216	**0.010**
Group: BD vs. control	0.253	0.017	1.034	0.315

COM	R^2	R^2 change	F for R^2 change	p
Model 1				
Rey-AVLT total	0.270	0.270	17.795	**<0.001**
Group: BD vs. control	0.312	0.042	2.872	0.097
Model 2				
VS reconstruction	0.092	0.092	4.864	**0.032**
Group: BD vs. control	0.181	0.089	5.081	**0.029**
Model 3				
Rey-AVLT total	0.270	0.270	17.795	**<0.001**
VS reconstruction	0.294	0.023	1.548	0.220
Group: BD vs. control	0.321	0.027	1.815	0.185
Model 4				
VS reconstruction	0.092	0.092	4.864	**0.032**
Rey-AVLT total	0.294	0.202	13.424	**0.001**
Group: BD vs. control	0.321	0.027	1.815	0.185

For the OLB measure, entry of the Rey-AVLT first explained 18.6% of the variance, with VSR not producing a significant increase (5.0%). However, even when entering VSR first, the subsequent addition of Rey-AVLT total produced a significant increase (11.7%). In the case of COM, the same pattern emerged, with entry of the Rey-AVLT first explaining 27.0% of the variance while the subsequent entry of VSR was not significant (2.3%). Entering VSR first explained 9.2% of the variance, with the subsequent entry of Rey-AVLT significantly increasing the proportion explained (20.2%). In both analyses of OLB and COM, the final entry of the group variable was not significant, explaining only an additional 1.7% and 2.7% respectively.

5.3.4 Is the effect group-specific?

In order to examine the specificity of this effect to the individual groups, a similar series of analyses was performed for each group independently.

Table 5-6. POM hierarchical regression (separate groups, with Rey-AVLT total and VSR)

	Controls				Patients			
POM	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
Model 3								
Rey-AVLT total	0.001	0.001	0.031	0.861	0.032	0.032	0.764	0.391
VS reconstruction	0.255	0.253	7.471	**0.012**	0.106	0.074	1.817	0.191
Model 4								
VS reconstruction	0.231	0.231	6.908	**0.015**	0.079	0.079	1.961	0.175
Rey-AVLT total	0.255	0.024	0.694	0.414	0.106	0.027	0.676	0.420

Table 5-7. OLB and COM hierarchical regression (separate groups, with Rey-AVLT total and VSR)

	Controls				Patients			
OLB	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
Model 3								
Rey-AVLT total	0.028	0.028	0.656	0.426	0.246	0.246	7.522	**0.012**
VS reconstruction	0.159	0.132	3.448	0.077	0.262	0.016	0.469	0.501
Model 4								
VS reconstruction	0.159	0.059	4.352	**0.048**	0.022	0.022	0.528	0.475
Rey-AVLT total	0.159	<0.001	0.010	0.922	0.262	0.240	7.148	**0.014**

	Controls				Patients			
COM	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
Model 3								
Rey-AVLT total	0.168	0.168	4.656	**0.042**	0.203	0.203	5.841	**0.024**
VS reconstruction	0.203	0.035	0.953	0.340	0.206	0.003	0.087	0.771
Model 4								
VS reconstruction	0.106	0.106	2.723	0.133	0.006	0.006	0.141	0.711
Rey-AVLT total	0.203	0.097	2.678	0.116	0.206	0.200	5.527	**0.028**

Significance of entry indicated by shading.

It is clear from the above analyses that a very consistent pattern emerges in the results of the hierarchical regression. In controls, VSR explained a significant proportion of the variance in POM, irrespective of the order of entry into the model. For OLB, VSR was also significant but only when entered first into the model, similarly for COM, although the Rey-AVLT explained a significant proportion of the variance in COM when entered first, this was not significant in model 4 when entered after VSR (explaining <10% additional variance).

In patients, the results were very different. For POM, neither variable was significant. However, for OLB and COM, the Rey-AVLT explained a significant proportion of the variance irrespective of the order of entry (≥20% for all) while the addition of VSR was non-significant for all (≤2.2% variance explained).

5.4 Discussion of Chapter 5

In this chapter, the results of the first use of the ORT paradigm in patients with bipolar depression and healthy controls are presented. Initial univariate analyses demonstrated large effect sizes on all three primary outcome measures from the task: OLB (d= -0.70), COM (d= -0.78) and POM (d= -1.02). One of the control tasks (VSR) also showed a between group difference. When the results of this task were covaried, only the group difference in POM remained significant. Analysis of the secondary neuropsychological tests revealed a broad pattern of performance impairment in the bipolar patients. Correlational analysis of the relationships between the secondary battery and ORT measures suggested a different pattern of significant relationships between patients and controls. Given the results of the earlier chapters it is of note for example that in controls POM correlates significantly with SWM between errors, while in patients it is COM which correlated most strongly with SWM between errors.

There was a significant impairment in verbal learning in the patient group (Rey-AVLT total; d= -0.79). In patients, performance on this measure correlated significantly with the OLB and COM measures, but not with POM (no equivalent significant relationships were found in controls). Subsequent hierarchical multiple regression revealed that all but a trivial proportion of variance in OLB and COM measures (over and above that explained by control variables i.e. VSR) can be explained by verbal learning with 'group' explaining only a trivial amount of variance. In the equivalent analysis of POM, the addition of 'group' still explained significant additional variance (~10%).

Similar to the effect reported in many of the previous studies using the ORT paradigm (see section 5.1), we see here an apparent separation between the POM process and the OLB/COM processes. Although this did not meet criteria for a differential deficit, the pattern

of correlation, subsequent ANCOVA and results of the multiple regression analyses shows an apparent difference in the mediating effect of verbal learning between patients and controls (this will be discussed in detail in sections 5.4.1 and 7.3 below).

5.4.1 Theoretical implications of the ORT findings

In a study specifically designed to examine the role of verbal memory processes and object-location memory, Kessels and Postma (2002) examined the effect of articulatory suppression (AS+; counting 1 to 5 recurrently) or silence (AS-) during either the encoding or maintenance phase of the three ORT processes. Performance was significantly worse with AS+ on the OLB and COM conditions, but there was no effect on POM. Also, the effect only occurred during the encoding phase; there was no effect on any measure when AS+ was applied during maintenance. These results were discussed in terms of the earlier findings of Postma and de Haan (1996), that the effect of AS+ appears to occur even with stimuli that are not readily nameable and therefore that it may not be a direct verbal effect *per se*, but possibly a disruption in underlying categorical information processing.

The implications of the above discussion can be outlined in a number of hypothesis, although it is again noted that these are purely speculative as they were not tested directly in the present study.

5.4.1.1 What underpins the link between some ORT measures and Rey-AVLT total learning?

It is first important to again consider precisely what traits or capacities underlie the verbal learning measure (Rey-AVLT total) and also the ORT measures, OLB and COM[26]. In terms of the former, there is a great deal of work in the hemispheric preponderance literature highlighting the close relationship between language processing and categorization or the ability to process information in a categorical manner (Kosslyn et al., 1989; Parrot et al., 1999). Therefore, whilst the Rey-AVLT is clearly a verbal memory measure, it may also be viewed to some degree as measure of categorical learning or processing. Furthermore, when referring to categorical relations of objects, linguistic and perceptual/cognitive representations of space are at least partially distinct (although it has been noted that language can modify both perceptual sensitivity and cognitive style- for a review see Kemmerer, 2006). For example, evidence from a seminal paper by Kemmerer and Tranel (2000) reports a double dissociation in two brain-damaged subjects between linguistic representation of spatial relationships and perceptual representations, especially those of a categorical nature. Jager & Postma (2003) summarise this as evidence of a "tripartion [*sic*] between perceptual-coordinate spatial codes, perceptual-categorical codes and verbal-categorical spatial codes".

Therefore, within the tripartition described by Jager and Postma above and the consideration of what is being measured by the Rey-AVLT total verbal learning measure (i.e. Rey-AVLT being foremost a verbal memory measure, but possibly also tapping some element of general categorical processing) a number of hypotheses can be considered to explain the pattern of results from the regression analyses. These are outlined below as individual potential hypotheses although it should be noted that they may not be mutually exclusive, but may in fact overlap.

[26] It is noted that COM is a complex measure which may be considered to include both categorical and coordinate processes. Here, for the purposes of the discussion, it is included along with OLB due to the correlation between OLB and COM in both groups, and a similar pattern/relationship of both measures to predictor variables in the regression analyses.

The first hypothesis is that the relationship between verbal learning and OLB/COM is because language processing in general overlaps (as regards the processes involved) the binding and relational aspect of memory for spatial arrays in that both can involve a form of categorical processing. This would explain the loss of the significant group effect in OLB/COM when factoring verbal learning into the statistical model. However, the proportion of the variance in OLB/COM measures explained by verbal learning was greater in patients than controls. Considering this, a second hypothesis could be that patients preferentially attempt to verbally encode items (*"the chair was top/left, the car was bottom/right ..."*), while controls perceptually encode items with either no or minimal support from other (verbal) processes. This would also explain the results of the regression analyses where a larger proportion of the variance of the group difference in OLB/COM was explained by verbal learning. In this context, when the groups are examined separately, the greater proportion of variance explained by verbal learning in patients is because the verbal learning measure has much greater loading to verbal-categorical processes. A third related hypothesis is that patients having visuospatial (perceptual) memory impairment, including an impairment of visuospatial metric processing, draw more on verbal/verbal-categorical processes to attempt to maintain or 'scaffold' performance on tasks which are amenable to such a strategy i.e. when objects are unique. However, performance on the verbal learning measure is statistically worse in patients (see section 5.3.2) therefore it is unclear why patients would use such a process, unless, despite being impaired, it is relatively 'less impaired' than other processes that could be used. This perhaps also explains the arithmetically larger effect size for the group difference in POM compared to OLB/COM.

In terms of the POM measure, memory for exact metric/coordinate locations is likely to be highly demanding of cognitive resources. Moreover, retention of exact spatial location is critically time-dependent with a rapid decay function. Some distortions in precise location are

evident after retention intervals in the order of hundreds of milliseconds or less (Werner & Diedrichsen, 2002) and these distortions increase over time (Postma et al., 2006). Recent work by van der Ham et al. (2007) examining the time-course of hemispheric specialisation for categorical and co-ordination spatial relations revealed that whilst the predicted *relative* hemispheric differences occurred at very brief retention intervals (500ms), the co-ordinate specialisation of the right hemisphere was not present when durations were 2 seconds or more (in fact there appeared to be a shift to LH advantage). It was reported that, subjectively, participants described the attempted use of verbal encoding strategies at longer intervals. The effect was summarised thus: "*it is very well possible that in the long interval, the coordinate strategy failed completely, because the exact, metric information had decayed in memory. A verbal approach could replace the coordinate strategy, which would result in a lower level of performance, because the coordinate trials were not perfectly solvable without knowing the total number of possible positions. A verbal, more categorical strategy, using words like 'near', 'in the middle', and 'far', could well have caused the left hemispheric advantage found, because of the use of categories and verbal memorization*".

5.4.1.2 Summary

Together the results of these studies suggest that it may be the rapid and accurate encoding of spatial configural information which may be at the heart of the performance on the ORT in the present study. It has been suggested that the formation of some categorical coding must occur very rapidly to maintain anything close to an accurate representation of the original locations (Werner & Diedrichsen, 2002). For the POM condition, the large effect size and statistical difference between groups – which remained after factoring in the variance explained by the control or verbal learning tasks – may result from an impairment in rapid encoding of precise spatial locations. Because all items are identical in this condition, as

discussed above, any switch to / or support from a more verbal strategy is unlikely to adequately scaffold performance.

These data raise an important empirical question for the final chapter in this thesis, namely, what is the interrelationship between spatial memory processes, broader neuropsychological components and the HPA axis? Also, do these processes operate differently in patients compared to healthy controls? Given the findings in this Chapter with regard to the differences in ORT measures and the contributions of other processes to task performance (e.g. the verbal memory findings), one primary aim is to ascertain if similar effects occur with the SWM task (i.e. are there differences in the processes underlying the between and within search errors). The aim of the subsequent chapter is to integrate the findings from the previous two chapters and introduce additional measures of HPA axis functioning which may be more specific or sensitive in order to examine these relationships.

Chapter VI

Neuropsychological and HPA axis correlates of spatial memory processes

6. Neuropsychological and HPA axis correlates of spatial memory processes

6.1 Introduction to Chapter 6

In this final chapter the relationships among spatial memory processes, neuropsychological composites and HPA axis measures will be examined in a series of multiple regression analyses. This will be completed in the overall sample of patients and controls as described in chapter 4 (n=100 total) and also in the sub-set of participants who completed the ORT paradigm as described in the preceding Chapter 5.

As discussed in the introduction to Chapter 1, there are many methods which can be used to assess HPA axis function in humans. Also, as identified in Chapters 2 and 3, to fully explore the role of the HPA axis as a mediator of neuropsychological performance, measures beyond peripheral cortisol secretion should be considered. Here, three different measures are included: salivary cortisol-DHEA ratio, the cortisol awakening response and the degree of dexamethasone non-suppression (using the DST). These three were selected to provide a more comprehensive assessment than any of the measures can individually, particularly as has been noted in earlier discussion, there are examples in patients with bipolar disorder where 'activating' tests (which more closely assess functional alterations at the receptor level) have demonstrated differences in HPA activity in the absence of basal salivary changes (e.g. Watson et al., 2004) and conversely, examples of instances where CAR abnormalities have been detected in the absence of abnormal DST responses (Deshauer et al., 2003). Of particular interest are the CAR and the DST which are more closely linked to the functioning of the GR – in the case of the CAR, because it assesses levels during the maximal waking surge and the DST, a GR agonist (see General introduction, section 1.4.1)

The neuropsychological measures used will be composite scores derived from the components extracted from the series of PCA analyses in Chapter 4. Initially, a general analysis will examine the relationship between the spatial memory measures and HPA axis measures. Subsequently, a more detailed analysis will be conducted with the SWM and the ORT task in order to examine the relationship of the neuropsychological composites to the different task processes. This analysis will also draw on individual variables of theoretical interest in order to examine their relationship to the visuo-spatial processes examined. With regard to the SWM, it is expected that a different pattern of loadings will be observed across the two error types: BSE and WSE. Based on the results of Chapter 5 from the Object Location Memory task, a specific confirmatory hypothesis can be put forward that along with visuospatial composites, (in patients) the verbal memory composite will predict significant variance in selected OLM measures – OLB and COM.

6.2 Methods

6.2.1 Subjects

Details of the subjects included in this chapter have been described previously (see section 4.2). General demographic and clinical information for this sample is also reported earlier (section 4.3).

6.2.2 HPA axis measures

For the assessment of HPA axis function, two saliva sampling methods and one plasma sampling method was adopted. Saliva sampling took place on separate days, prior to completion of neuropsychological testing. Due to the administration of dexamethasone during the plasma sampling protocol, this was completed after neuropsychological testing.

The first saliva sampling method was 8am/8pm cortisol, DHEA, and cortisol-DHEA (CD) ratio. Saliva was collected by a passive drool method at the two time points over a single day. Samples were analysed to measure cortisol, DHEA and the CD ratio. The second saliva method utilised was the cortisol awakening response (CAR). Five saliva samples were taken (again using the passive drool method) from the time of awakening, then at 15-minute intervals thereafter for one hour. All saliva samples were assayed for cortisol and DHEA. Details of the assay procedures are presented previously (section 2.2.5).

For the dexamethasone (dex) suppression test (DST), a blood sample (5ml) was taken after neuropsychological testing to serve as a baseline. One mg dex was taken orally at 11pm the same evening and the participant returned for a repeat blood sample the next day at the same time as the baseline had been assessed. The blood sample was assayed for cortisol and the extent of suppression (DST Δ) assessed by subtracting the post-dex level from the pre-dex level.

6.2.3 Statistical analysis

Due to the iterative nature of this chapter, specific details of each analysis will be given throughout the results section. The general approach adopted will be to utilised multiple regression with forced entry. The principal steps of interest will be the R^2 change for each variable/composite entry within the model. Where tests have relevant control tasks associated with them (for example, the VSR of the ORT) in exploratory models these variables will be entered first in order to examine the additional variance explained by the true variables of interest. In general, one of the main aims of these analyses will be to examine the contribution of verbal memory measures to the performance of the spatial tasks described in the previous chapters of the thesis. Also, although the initial analyses of the separate groups will all be performed using the control-derived PCA composites, the analyses will be repeated

in the patient group using their own PCA derived composites in order to examine any differences. In effect, this will establish if any absence of significant relationships between the composite predictors and dependent variable are true and not a consequence of the underlying neuropsychological variable loadings being different in the patient group and therefore that composite being unrepresentative of that given factor in the patients).

6.3 Results

Presented first is a brief summary of the results of the additional HPA axis assessments.

6.3.1 Additional neuroendocrine measures

6.3.1.1 Data summary

Of the total sample of n=100, not all participants completed the HPA axis assessments. Due to the issue of artificially reducing variance when large numbers of data points are imputed, the analysis was carried out on observed data only. The only exception was that for the CAR, where a data point was missing in between two valid observations, the missing point was imputed as the midpoint between the two. Where the first or last sample was missing, it was replaced using the group mean. Replacement of this type was only performed where more than half the data was present i.e. for only those participants with 2 or fewer missing observations.

For the 8am/8pm cortisol and DHEA data, full datasets were available for 27 patients (51%) and 35 controls (74%). For the remaining 26 patients: 16 had provided no samples, 1 had a missing 8am sample, 3 had 8pm cortisol sample missing, 2 had 8am DHEA missing, 2 had 8pm DHEA missing, 1 had cortisol analysed but no DHEA available, and 1 had the 8pm cortisol and

all DHEA missing. For the remaining 12 controls: 7 had provided no saliva samples, 3 had cortisol analysed but no DHEA was available, 1 had a the 8am cortisol sample missing, and 1 the 8pm DHEA sample missing. The reasons for missing data are varied, however instances where samples are provided but measures of either DHEA or cortisol are missing are generally the result of insufficient sample remaining to complete both assays.

For the CAR, full data sets were available for 36 patients (68%) and 34 controls (72%). Of the remaining patients, 1 had provided no samples, 8 had single missing samples (0mins, four at 30mins, two at 45mins, 60mins), 5 had two missing samples (0mins, 15mins / 0mins, 45mins / 15mins, 45mins / 30mins, 45mins / 30mins, 60mins), 1 had three missing samples (15mins, 45mins, 60mins), and two had all samples after 0mins missing. Of the remaining 13 controls: 8 had provided no samples, 2 had single missing samples (0mins/45mins), 1 had two missing samples (0mins, 45mins), and 2 had three missing samples (0mins, 15mins, 60mins / 30mins, 45mins, 60mins). Using the method of handling missing data described above resulted in the data of an additional 3 controls and 13 patients being available giving usable data for 49 (92%) patients and 37 (79%) controls.

For the DST, 29 patients (55%) and 30 controls (64%) provided full data sets. Of the remainder, only 5 patients and 10 controls did not provide any blood samples. Seventeen patients and 6 controls provided pre-dex samples but no post-dex sample, while 2 patients and 1 control had only the post-dex sample available. The main reason for instances where the pre-dex sample only was available was because a more detailed blood analysis was performed for the overall program of research of which this study was part, therefore participants agreed to this sample but either declined to take dexamethasone or could not attend the following day to provide an additional sample, so dexamethasone was not dispensed.

6.3.1.2 Analysis of neuroendocrine data

Salivary cortisol and DHEA data were examined separately in a repeated measures ANOVA with time as the within subjects factor. For cortisol, there was a significant effect of time (F=67.326, df=1,66, p<0.0001) but no significant main effect of group (F=0.109, df=1,66, p=0.743) or time by group interaction (F=0.240, df=1,66, p=0.626). Similarly, for DHEA, there was a significant effect of time (F=51.067, df=1,67, p<0.0001) but no significant main effect of group (F=0.277, df=1,67, p=0.277) or time by group interaction (F=0.077, df=1,67, p=0.782). Consequently, there was no difference between the groups in the molar cortisol-DHEA ratio.

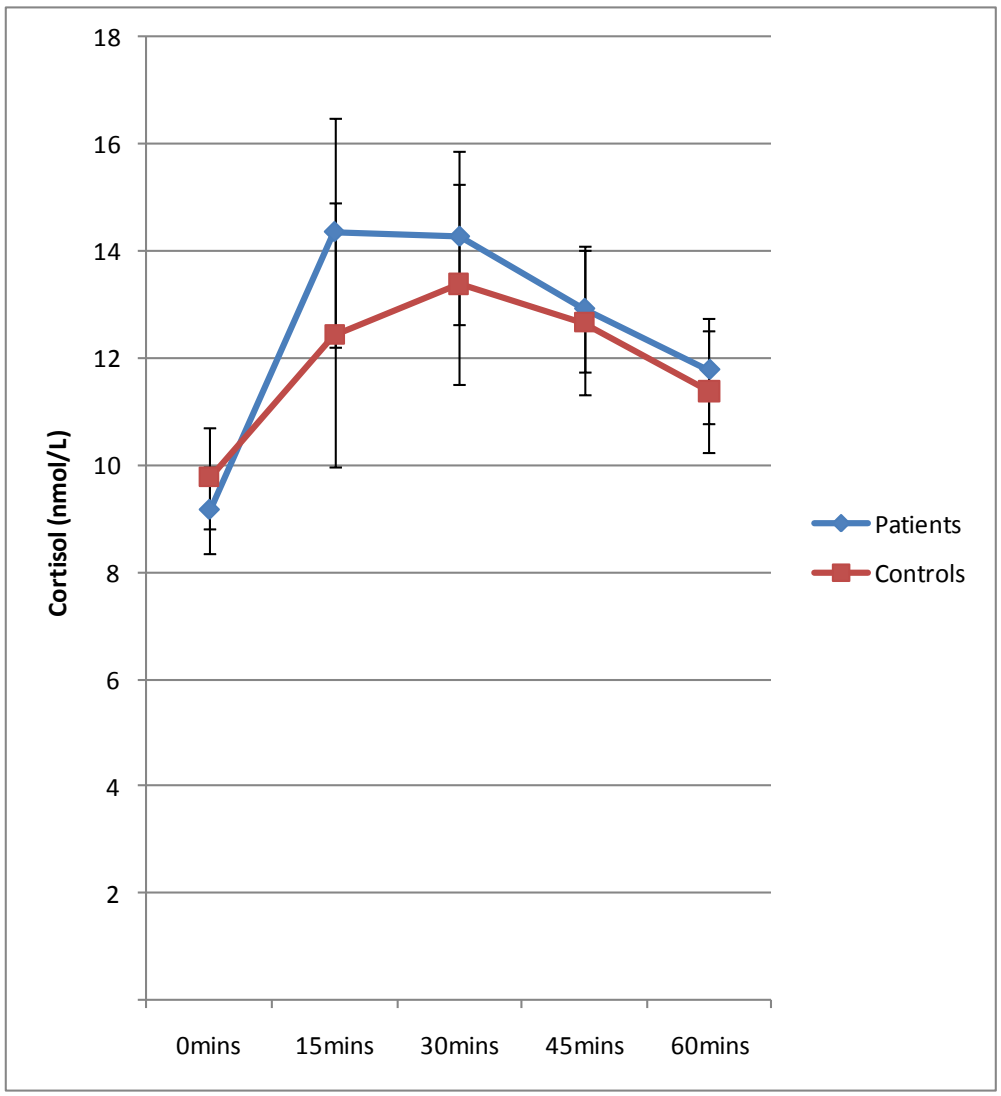
Table 6-1. Summary data for the endocrine measures.

	Controls		**Patients**			
	mean	s.d.	mean	s.d.	t	p
Cortisol-DHEA[a]						
8am cortisol	11.0	6.3	11.3	8.5		
8pm cortisol	2.9	5.4	2.2	1.7		
8am DHEA	4.5	4.1	3.9	2.9		
8pm DHEA	1.7	1.3	1.3	0.9		
8am CD ratio	4.3	3.9	4.3	4.4	0.041	0.986
8pm CD ratio	2.1	4.4	2.7	3.3	1.000	0.321
CAR[a]						
AUC	735.9	374.8	779.7	636.3	0.373	0.710
DST[a]						
Pre-dex cortisol	266.0	85.9	247.7	107.2		
Post-dex cortisol	46.2	102.5	51.0	54.4		
Δ cortisol	-229.5	80.6	-199.8	106.3	1.212	0.230

[a] *All values in nmol/L or nmol/L/min for AUC.*

Individual time points for the CAR are presented in Figure 6-1 below. Again there was a significant effect of time (F=5.545, df=4,336, p=0.006) but no significant main effect of group (F=0.106, df=1,84, p=0.745) or time by group interaction (F=0.389, df=4,36, p=0.655).

Figure 6-1. The cortisol awakening response (CAR)



Finally, for the DST, as can be seen in Table 6-2 there was no significant difference in the degree of cortisol suppression between the groups following dexamethasone. However, the DST typically uses a cut-off to define non-suppression and reduces the data to categorical values (i.e. suppressor vs. non-suppressor). The historical value based on the original paper (Carroll et al., 1981) was 5µg/dl (138nmol/L) although this value was based on the optimum cut-off for sensitivity and specificity. Therefore an ROC analysis was performed on the Δ

cortisol values in patients and controls. The AUC of the ROC was 0.654 (95%CI=0.502 to 0.806). An optimum cut-point of -178.9 nmol/L produced a sensitivity of 0.55 and specificity of 0.77. Examination of the cross-tabulation of patient and controls by DST suppression status was significant (χ^2=6.284, *df*=1, *p*=0.017).

Table 6-2. Number of patients and controls classified as DST non-suppressors

	Patients	Controls
Non-suppressor	16 (55.2%)	7 (23.3%)
Suppressor	13 (44.8%)	23 (76.7%)

Percentage with reference to proportions within diagnostic group.

6.3.2 Neuroendocrine measures as predictors of spatial memory

From the additional HPA axis measures, four principal outcome measures were examined in relation to the spatial memory tasks: 8am CD ratio, 8pm CD ratio, CAR AUC, DST Δ cortisol. The spatial memory outcomes were restricted to SWM between and within search errors, and Object Relocation Test POM, OLB and COM.

Table 6-3. Correlation between spatial memory measures and HPA axis measures

Whole group	**SWM**		**ORT**		
	Between	Within	POM	OLB	COM
8am CD ratio	-0.029	0.122	0.045	**-0.415***	-0.183
8pm CD ratio	-0.137	-0.059	0.204	-0.187	-0.159
CAR AUC	-0.019	0.035	0.200	-0.090	0.151
DST Δ cortisol	-0.036	-0.173	-0.117	-0.047	-0.020

**p<0.05*

	SWM		ORT		
	Between	Within	POM	OLB	COM
Patients					
8am CD ratio	-0.091	0.086	-0.156	-0.496	-0.358
8pm CD ratio	-0.133	0.064	0.253	-0.086	-0.187
CAR AUC	0.103	-0.055	0.166	0.060	**0.422***
DST Δ cortisol	-0.060	-0.254	-0.293	-0.442	-0.393
Controls					
8am CD ratio	0.038	0.154	0.319	**-0.507***	-0.114
8pm CD ratio	-0.134	-0.192	0.232	-0.347	-0.082
CAR AUC	-0.135	0.162	0.086	**-0.490***	-0.228
DST Δ cortisol	-0.008	0.006	-0.400	0.221	-0.059

**p<0.05*

As can be seen in the table, there is little evidence of any consistent relationship between HPA measures and spatial tests. Between the groups, the only significant correlations were for the CAR to be associated with poorer COM performance in patients and for both CAR and the 8am CD ratio to be associated with better OLB performance in controls. Individual composites were also examined but again there were no meaningful significant correlations for any measure (see Appendix 9.6, p.376).

As the only measure to show significance in the whole group was CAR AUC for the OLB measure, a regression was performed with entry of this measure followed by group. From the analysis below the initial entry adds only 3% of variance. However subsequent entry of group adds significant additional variance (~10%).

	R^2	R^2 change	F for R^2 change	p
OLB				
CAR AUC	0.003	0.003	0.141	0.709
Group: BD vs. control	0.106	**0.106**	5.211	**0.027**

Finally, to examine the effect of DST suppression status, independent t-tests were used to examine for differences in spatial memory processes between suppressors and non-suppressors (diagnostic group was not considered in this analysis). There were no significant differences between groups (suppression status) for SWM BSE (t=0.641, df=57, p=0.524), WSE (t=0.443, df=57, p=0.660), ORT POM (t=0.029, df=28, p=0.977), OLB (t=0.054, df=28 p=0.957) or COM (t=0.120, df=28, p=0.906). There were also no significant differences in any of the composite measures (p>0.1; data not shown).

6.3.3 Neuropsychological composites as predictors of spatial memory performance

Table 6-4. Mean composite scores (derived from individual variable z scores) for each component of the PCA analysis

	Bipolar (n=53)			Controls (n=47)				
	Mean	s.d.	SEM	Mean	s.d.	SEM	t	p
Optimised model							*(df=98)*	
TOTAL_c1 (visuo-spatial; VS) [a]	-0.205	0.735	0.101	0.229	0.762	0.111	-2.900	**0.005**
TOTAL_c2 (verbal memory) [b]	-0.297	0.847	0.116	0.335	0.796	0.116	-3.828	**<0.001**
TOTAL_c3 (verbal exec) [c]	-0.298	0.649	0.089	0.336	0.833	0.122	-4.274	**<0.0001**
control_c1 (verbal memory) [b]	-0.297	0.847	0.116	0.335	0.796	0.116	-3.828	**<0.001**
control_c2 (VS complex) [d]	-0.165	0.818	0.112	0.186	0.851	0.124	-2.098	**0.038**
control_c3 (VS immediate) [e]	-0.268	0.713	0.098	0.301	0.781	0.114	-3.808	**<0.001**
control_c4 (verbal exec) [f]	-0.357	0.724	0.099	0.402	0.928	0.135	-4.587	**<0.0001**
patient_c1 (VS) [g]	-0.220	0.698	0.096	0.247	0.660	0.096	-3.424	**0.001**
patient_c2 (verbal memory) [b]	-0.297	0.847	0.116	0.335	0.796	0.116	-3.828	**<0.001**
patient_c3 (digit srec) [h]	-0.162	0.817	0.112	0.182	0.774	0.113	-2.153	**0.034**
Optimised SWM model								
TOTAL_c1 (verbal memory) [b]	-0.297	0.847	0.116	0.335	0.796	0.116	-3.828	**<0.001**
TOTAL_c2 (VS) [i]	-0.238	0.740	0.102	0.266	0.835	0.122	-3.200	**0.002**
TOTAL_c3 (verbal exec) [c]	-0.298	0.649	0.089	0.336	0.833	0.122	-4.274	**<0.0001**
control_c1 (verbal memory) [b]	-0.297	0.847	0.116	0.335	0.796	0.116	-3.828	**<0.001**
control_c2 (verbal exec) [c]	-0.298	0.649	0.089	0.336	0.833	0.122	-4.274	**<0.0001**
control_c3 (VS) [i]	-0.238	0.740	0.102	0.266	0.835	0.122	-3.200	**0.002**
patient_c1 (VS) [j]	-0.253	0.688	0.094	0.283	0.676	0.099	-3.915	**<0.001**
patient_c2 (verbal memory) [b]	-0.297	0.847	0.116	0.335	0.796	0.116	-3.828	**<0.001**
patient_c3 (digit srec) [h]	-0.162	0.817	0.112	0.182	0.774	0.113	-2.153	**0.034**

Key to z-score variables that constitute each composite:

[a] SWM between errors, strategy; Spatial span forward, reverse; VPT.

[b] Rey-AVLT total, A7, recognition.

[c] FAS, ELFT, Digit Span reverse.

[d] SWM between errors, strategy; SOPT.

[e] Spatial span forward, reverse; DSST

[f] FAS; ELFT

[g] SWM between errors, strategy; Spatial span forward, reverse; VPT; FAS.

[h] Digit span reverse; SRec

[i] Spatial span forward, reverse; VPT.

[j] Spatial span forward, reverse; VPT; FAS.

Table 6-4 above summarises the composite scores for each of the PCA analyses from Chapter 4. As can be seen, all composites indicate significantly worse scores in patients compared to controls. All values are included although it should be noted that there are several repetitions included in the table (e.g. the verbal memory composite) where several models produced the same solutions in terms of the individual variables constituting each.

Given the findings in Chapter 3 of specific changes in SWM (BSE not WSE), these two error types are examined first in order to explore any differences in underlying composite loadings.

6.3.3.1 Spatial Working Memory

6.3.3.1.1 Between search errors

In order to restrict the number of contrasts performed, an initial analysis was performed on the whole dataset (n=100) with concurrent forced entry of the three total composite measures (from the Optimised SWM model). The dependent variable was square root transformed prior to analysis.

The overall model was highly statistically significant (F=31.02, df=3,96, p<0.0001) with 49.2% of the variance explained (R=0.702). Examination of the individual composite coefficients revealed that the verbal executive composite (TOTAL_c3) was not statistically significant (B=-0.038, standard error, SE=0.229; β=-0.13; t=-0.165, p=0.869), therefore only the verbal memory (c1) and visuospatial memory (c2) composites were used further in the hierarchical models.

For each model, group (patient or control) was entered after the entry of the composites to assess the additional variance explained. Model 1 entered the c1 then c2, while in Model 2 the

order was reversed. For the remaining analyses, the group variable was removed and patients any controls were examined separately. The first set of models used the composites derived from the healthy controls' PCA analysis (SWM optimised; control c1 and c3). The last set of models examined patients only, using the composites derived from their own PCA analysis (SWM optimised; patients c1 and c2).

Table 6-5. SWM between search errors (TOTAL group)

	R^2	R^2 change	F for R^2 change	p
Model 1				
TOTAL c1 (verbal)	0.232	**0.232**	29.656	**<0.001**
TOTAL c2 (VS)	0.492	**0.260**	49.579	**<0.001**
Group: BD vs. control	0.505	0.013	2.467	0.120
Model 2				
TOTAL c2 (VS)	0.418	**0.418**	70.389	**<0.001**
TOTAL c1 (verbal)	0.492	**0.074**	14.135	**<0.001**
Group: BD vs. control	0.505	0.013	2.467	0.120

Table 6-6. SWM between search errors (groups separated; control derived composites)

	Controls				**Patients**			
	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
Model 3								
control c1 (verbal)	0.208	**0.208**	11.816	**0.001**	0.214	**0.214**	13.914	**<0.001**
control c3 (VS)	0.467	**0.259**	21.351	**<0.001**	0.517	**0.302**	31.283	**<0.001**
Model 4								
control c3 (VS)	0.390	**0.390**	28.755	**<0.001**	0.415	**0.415**	36.241	**<0.001**
control c1 (verbal)	0.467	**0.077**	6.342	**0.015**	0.517	**0.101**	10.481	**0.002**
Model 5								
patient c2 (verbal)	-	-	-	-	0.214	**0.214**	13.914	**<0.001**
patient c1 (VS)	-	-	-	-	0.481	**0.267**	25.751	**<0.001**
Model 6								
patient c1 (VS)	-	-	-	-	0.382	**0.382**	31.566	**<0.001**
patient c2 (verbal)	-	-	-	-	0.481	**0.099**	9.556	**0.003**

As can be seen in all of the above models, irrespective of the order of entry, additional variance was explained by both the visuospatial and verbal composites. This occurred for the group as a whole and for patients and controls separately. Overall it appeared that arithmetically the visuospatial composite may be the greater predictor as R^2 values were always higher and when it was entered first (model 4 and 6) around 40% of the variance was explained by this variable. Nevertheless, entry of the verbal composite always explained significant additional variance (around 20% when entered first, and around 10% when entered second).

6.3.3.1.2 Within search errors

The same analyses were also performed (Models 1 to 6) for the SWM within search error rate. Again, the square root transformed values were utilised (however, it should be noted that the high number of zero error scores on this variable meant that while the distribution was improved, it did not achieve normality).

Table 6-7. SWM within search errors (TOTAL group)

	R^2	R^2 change	F for R^2 change	p
Model 1				
TOTAL c1 (verbal)	0.057	**0.057**	5.967	**0.016**
TOTAL c2 (VS)	0.116	**0.059**	6.468	**0.013**
Group: BD vs. control	0.143	0.027	3.009	0.086
Model 2				
TOTAL c2 (VS)	0.097	**0.097**	10.528	**0.002**
TOTAL c1 (verbal)	0.116	0.019	2.119	0.149
Group: BD vs. control	0.143	0.027	3.009	0.086

Table 6-8. SWM within search errors (groups separated; control derived composites)

	Controls				Patients			
	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
Model 3								
control c1 (verbal)	0.123	**0.123**	6.318	**0.016**	0.044	0.044	2.373	0.130
control c3 (VS)	0.248	**0.125**	7.319	**0.010**	0.091	0.047	2.576	0.115
Model 4								
control c3 (VS)	0.198	**0.198**	11.081	**0.002**	0.068	0.068	3.711	0.060
control c1 (verbal)	0.248	0.051	2.961	0.092	0.091	0.023	1.290	0.262
Model 5								
patient c2 (verbal)	-	-	-	-	0.044	0.044	2.373	0.130
patient c1 (VS)	-	-	-	-	0.117	**0.073**	4.132	**0.047**
Model 6								
patient c1 (VS)	-	-	-	-	0.099	**0.099**	5.623	**0.022**
patient c2 (verbal)	-	-	-	-	0.117	0.018	1.024	0.316

In comparison to the between search errors, the regression models on within search errors indicated minimal involvement of the verbal composite. In the total sample and in controls, while initial entry of verbal memory was significant, when the visuospatial composite was entered first no significant additional variance was explained by the entry of the verbal composite (models 1 to 4). Interestingly, in patients, when the composites derived from the control participants PCA solutions were used, entry of neither composite was significant. However, when composites derived from their own PCA solution were used, an even clearer pattern was observed with verbal memory explaining no significant additional variance, irrespective of the order of entry. Only the visuospatial composite explained significant variance (models 5 and 6).

6.3.3.2 Object Relocation Task

6.3.3.2.1 Whole sample

An initial analysis was performed on the whole sample who had completed the ORT (n=50). Due to the number of variables (with the inclusion of group), the hierarchical regression was simply performed using entry of the optimised composites in order. This order reflects the overall loadings from the original PCA analysis in chapter 4 i.e. the relative contribution of each component to the overall PCA model. All components were used as the aim was to establish not just loadings of each but to examine the effect of final entry of 'group'.

Table 6-9. ORT task (TOTAL group)

	R^2	R^2 change	F for R^2 change	p
POM				
TOTAL c1 (VS)	0.205	**0.205**	12.396	**0.001**
TOTAL c2 (verbal)	0.212	0.007	0.422	0.519
TOTAL c3 (verbal exec)	0.218	0.006	0.336	0.565
Group: BD vs. control	0.377	**0.159**	11.460	**0.001**
OLB				
TOTAL c1 (VS)	0.063	0.063	3.228	0.079
TOTAL c2 (verbal)	0.206	**0.143**	8.489	**0.005**
TOTAL c3 (verbal exec)	0.252	0.045	2.776	0.103
Group: BD vs. control	0.276	0.025	1.530	0.222
COM				
TOTAL c1 (VS)	0.126	**0.126**	6.918	**0.011**
TOTAL c2 (verbal)	0.364	**0.138**	8.777	**0.005**
TOTAL c3 (verbal exec)	0.282	0.018	1.162	0.287
Group: BD vs. control	0.316	0.034	2.264	0.139

While aspects of this initial analysis replicate what was established in chapter 5, it is interesting to note that the pattern of the variance explained by the composites is different for POM than for OLB and COM. POM is the only process for which the final entry of group

into the model explains significant additional variance (~16%) compared with OLB and COM (~3%). Also, while the initial entry of the visuospatial composite is significant in all 3 models (trend with OLB), subsequent entry of the verbal composite is only significant for OLB and COM.

The remaining analyses are all performed on the groups separately. Following a similar procedure as above, the regression models will use the control-derived composite scores and then the patients models will be re-calculated using patient-derived scores. Due to the number of permutations that are possible with the entry of five variables in a hierarchical fashion, the models will be constructed systematically to focus primarily on the visuospatial versus verbal contribution to each of the three ORT variables. Later models in each section will begin with the inclusion of the visuo-spatial reconstruction (VSR) measure to remove this source of variance (see section 5.3.1) and then examine the additional unique contribution of verbal memory.

6.3.3.2.2 Position-only memory (POM)

Based on the whole sample analysis, the visuospatial (VS) measures are of particular interest in these initial POM analyses. Both control derived VS composites (optimised model: c2 and c3; a complex visuospatial measure, consisting of SWM between search errors and strategy and the SOPT, and a more immediate one[27], consisting of spatial span forward and reverse and the DSST), were entered first in the initial model, followed by the remaining composites (note, the visuospatial and verbal composites were always sequential to each other, in order to allow for the maximum chance of additional variance to be explained by one or the other). Then this was repeated with the VS measures last.

[27] Note that these are fairly general identifying labels only; the individual variables that constitute each composite were in some instances difficult to classify under a single term. These will be discussed in greater detail in section 6.4.

Table 6-10. POM (groups separated; control derived composites)

	Controls				Patients			
	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
Model 1								
control c2|c3 (both VS)	0.349	**0.349**	5.893	**0.009**	0.285	**0.285**	4.376	**0.025**
control c1 (verbal)	0.404	0.055	1.937	0.179	0.350	0.065	2.106	0.161
control c4 (verbal exec)	0.406	0.002	0.075	0.786	0.427	0.077	2.698	0.116
Model 2								
control c1 (verbal)	0.018	0.018	0.416	0.525	0.024	0.024	0.565	0.460
control c4 (verbal exec)	0.046	0.028	0.654	0.427	0.061	0.037	0.874	0.260
control c2|c3 (both VS)	0.406	**0.360**	6.061	**0.009**	0.427	**0.366**	6.390	**0.007**

Table 6-11. POM (patients only; patient- derived composites)

	Controls				Patients			
	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
Model 1								
patient c1 (VS)	-	-	-	-	0.002	0.002	0.040	0.844
patient c2 (verbal)	-	-	-	-	0.024	0.022	0.501	0.486
patient c3 (digit srec)	-	-	-	-	0.025	0.001	0.021	0.885
Model 2								
patient c2 (verbal)	-	-	-	-	0.024	0.024	0.565	0.460
patient c1 (VS)	-	-	-	-	0.024	<0.001	<0.001	0.999
patient c3 (digit srec)	-	-	-	-	0.025	0.001	0.021	0.885

From the above models it is clear that the only composites that predict a significant proportion of the variance in POM are the visuospatial measures. It is of note that even entry at the end of the model explains around 36% in both patients and controls. Interestingly, when re-examining the patient group using the composite scores derived from their own PCA, there were no significant entry steps in any model. This may suggest that as a consequence of the 'blurred boundaries' of the composites in patients (i.e. the mixture of verbal and

visuospatial variables, as discussed in section 5.4) these have poorer predictive ability than the 'purer' control-derived measures which follow more circumscribed theoretical divisions.

The final analysis conducted was in order to examine the effect of initially entering the VSR control task variable into the model followed by the remaining control derived composites, with the visuospatial measures entered last in order to assess the additional variance explained.

Table 6-12. POM (groups separated; control derived composites with initial VSR entry)

	Controls				Patients			
	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
Model 1								
VSR	0.231	**0.231**	6.908	**0.015**	0.079	0.079	1.961	0.175
control c1 (verbal)	0.240	0.009	0.262	0.614	0.089	0.010	0.250	0.622
control c4 (verbal exec)	0.257	0.017	0.470	0.501	0.138	0.049	1.203	0.285
control c2|c3 (both VS)	0.457	0.200	3.505	0.051	0.439	**0.300**	5.080	**0.017**

From this analysis it is clear that even with prior entry of the VSR and all remaining composites, the final entry of the visuospatial composites adds a substantial proportion of additional variance (30% for patients and 20% for controls, although it is noted that this marginally fails to reach conventional significance in controls).

6.3.3.2.3 Object-location binding (OLB)

As with the previous analysis of POM, the initial models compare the hierarchical entry of the control derived PCA composites. Due to the emergence of a significant entry of the verbal executive composite in one of the models, one additional model was included with initial entry of this measure.

Table 6-13. OLB (groups separated; control derived composites)

	Controls				Patients			
	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
Model 1								
control c2|c3 (both VS)	0.173	0.173	2.308	0.123	0.077	0.077	0.917	0.414
control c1 (verbal)	0.190	0.017	0.433	0.518	0.282	**0.206**	6.015	**0.023**
control c4 (verbal exec)	0.324	0.134	3.955	0.061	0.364	0.082	2.571	0.124
Model 2								
control c1 (verbal)	0.083	0.083	2.077	0.163	0.231	**0.231**	6.913	**0.015**
control c4 (verbal exec)	0.246	**0.163**	4.761	**0.040**	0.263	0.032	0.955	0.339
control c2|c3 (both VS)	0.324	0.078	1.152	0.336	0.364	0.101	1.590	0.229
Model 3								
control c4 (verbal exec)	0.195	**0.195**	5.572	**0.027**	0.008	0.008	0.183	0.673
control c2|c3 (both VS)	0.320	0.125	1.938	0.169	0.184	0.176	2.268	0.128
control c1 (verbal)	0.324	0.003	0.101	0.754	0.364	**0.180**	5.665	**0.027**

From the above analysis it is clear that there are consistent but different key composite loadings between the groups – the verbal executive composite (consisting of the verbal fluency measures) in controls and the verbal memory composite (consisting of Rey-AVLT total and delayed recall and recognition) in patients. In the case of the patients, even entering the verbal composite last in the model contributes to a significant increase in variance explained (~18%). In the case of the controls, although final entry of the verbal executive measure just fails to reach statistical significance, an additional 13.4% of the variance is explained, and in other models initial or second entry positions are significant (n.b. model 2 above was repeated with c2|c3 first and this second entry step of c4 is also significant; R^2_{change}=14.7%).

The next series of models was performed in order to repeat the analysis in patients using composites derived from their own PCA analysis. As can be seen in the table below, the profile of results was identical to that achieved using the control derived composites, with the entry of the verbal memory composite adding at least 21%, irrespective of the position within the model.

Table 6-14. OLB (patients only; patient- derived composites)

	Controls				Patients			
	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
Model 1								
patient c1 (VS)	-	-	-	-	0.018	0.018	0.410	0.528
patient c2 (verbal)	-	-	-	-	0.231	**0.214**	6.112	**0.022**
patient c3 (digit srec)	-	-	-	-	0.232	0.001	0.019	0.891
Model 2								
patient c2 (verbal)	-	-	-	-	0.231	**0.231**	6.913	**0.015**
patient c3 (digit srec)	-	-	-	-	0.232	0.001	0.021	0.887
patient c1 (VS)	-	-	-	-	0.232	<0.001	<0.001	0.997
Model 3								
patient c3 (digit srec)	-	-	-	-	0.006	0.006	0.129	0.723
patient c1 (VS)	-	-	-	-	0.020	0.014	0.319	0.578
patient c2 (verbal)	-	-	-	-	0.232	**0.212**	5.797	**0.025**

The final analysis conducted was in order to examine the effect of initially entering the VSR control task variable into the model followed by the remaining control derived composites. From the previous analyses it was identified that the verbal executive was the strongest predictor in control participants, while it was the verbal memory composite in patients. Therefore, these variables were entered last in order to assess the additional variance explained.

Table 6-15. OLB (groups separated; control derived composites with initial VSR entry)

	Controls				Patients			
	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
Model 1								
VSR	0.159	**0.159**	4.352	**0.048**	0.022	0.022	0.528	0.475
control c2|c3 (both VS)	0.231	0.072	0.987	0.389	0.082	0.060	0.684	0.516
control c4 (verbal exec)	0.385	0.153	4.981	0.037	0.189	0.107	2.640	0.120
control c1 (verbal)	0.385	0.001	0.020	0.889	0.364	0.175	5.231	**0.034**
Model 2								
VSR	0.159	**0.159**	4.352	**0.048**	0.022	0.022	0.528	0.475
control c2|c3 (both VS)	0.231	0.072	0.987	0.389	0.082	0.060	0.684	0.516
control c1 (verbal)	0.241	0.010	0.266	0.612	0.283	**0.200**	5.587	**0.028**
control c4 (verbal exec)	0.385	0.144	4.445	0.049	0.364	0.082	2.440	0.138

From this analysis it is clear that even with prior entry of the VSR and all remaining composites, the final entry of the verbal executive composite in the case of controls or the verbal memory composite in the case of patients adds a substantial proportion of additional variance explained.

6.3.3.2.4 Combined condition (COM)

Finally, as with OLB and COM, the initial models compare the hierarchical entry of the control derived PCA composites. For comparison, these are presented as they were with the OLB analysis.

Table 6-16. COM (groups separated; control derived composites)

	Controls				Patients			
	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
Model 1								
control c2|c3 (both VS)	0.213	0.213	2.981	0.071	0.081	0.081	0.975	0.393
control c1 (verbal)	0.293	0.080	2.366	0.139	0.149	0.068	1.670	0.210
control c4 (verbal exec)	0.303	0.010	0.292	0.595	0.149	<0.001	<0.001	0.998
Model 2								
control c1 (verbal)	0.250	**0.250**	7.657	**0.011**	0.106	0.106	2.732	0.112
control c4 (verbal exec)	0.258	0.008	0.233	0.634	0.117	0.010	0.257	0.617
control c2|c3 (both VS)	0.303	0.045	0.652	0.532	0.149	0.033	0.383	0.687
Model 3								
control c4 (verbal exec)	0.026	0.026	0.609	0.443	0.025	0.025	0.596	0.488
control c2|c3 (both VS)	0.236	0.211	2.895	0.078	0.082	0.057	0.649	0.533
control c1 (verbal)	0.303	0.067	1.916	0.182	0.149	0.067	1.577	0.224

From the models above it is apparent that overall the pattern of results with COM is less clear-cut. In patients, there are no steps in the models which add significant variance. In controls, although the entry of the verbal composite in model 2 is the only significant effect, it should be noted that the combined entry of the VS adds almost 15% variance in model one, and adds an additional 21% in model three, although both just fail to reach significance.

The next series of models was performed in order to repeat the analysis in patients using composites derived from their own PCA analysis. Again, as can be seen in the table below, the profile of results was similar with respect to the control derived composites, with no steps adding significant variance.

Table 6-17. COM (patients only; patient- derived composites)

	Controls				Patients			
	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
Model 1								
patient c1 (VS)	-	-	-	-	0.047	0.047	1.125	0.300
patient c2 (verbal)	-	-	-	-	0.124	0.078	1.947	0.177
patient c3 (digit srec)	-	-	-	-	0.133	0.008	0.206	0.655
Model 2								
patient c2 (verbal)	-	-	-	-	0.106	0.106	2.732	0.112
patient c3 (digit srec)	-	-	-	-	0.110	0.004	0.102	0.752
patient c1 (VS)	-	-	-	-	0.133	0.022	0.541	0.470
Model 3								
patient c3 (digit srec)	-	-	-	-	0.001	0.001	0.023	0.881
patient c1 (VS)	-	-	-	-	0.053	0.052	1.207	0.284
patient c2 (verbal)	-	-	-	-	0.133	0.080	1.930	0.179
Model 4								
patient c2 (verbal)	-	-	-	-	0.106	0.106	2.732	0.112
patient c1 (VS)	-	-	-	-	0.124	0.018	0.451	0.509
patient c3 (digit srec)	-	-	-	-	0.133	0.008	0.206	0.655

The final analysis conducted was in order to examine the effect of initially entering the VSR control task variable into the model followed by the remaining control derived composites.

Table 6-18. COM (groups separated; control derived composites with initial VSR entry)

	Controls				Patients			
	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
Model 1								
VSR	0.106	0.106	2.723	0.113	0.006	0.006	0.141	0.711
control c2|c3 (both VS)	0.215	0.109	1.458	0.255	0.081	0.075	0.861	0.437
control c4 (verbal exec)	0.238	0.024	0.618	0.441	0.082	0.001	0.012	0.912
control c1 (verbal)	0.303	0.065	1.768	0.199	0.151	0.069	1.554	0.228
Model 2								
VSR	0.106	0.106	2.723	0.113	0.006	0.006	0.141	0.711
control c2|c3 (both VS)	0.215	0.109	1.458	0.255	0.081	0.075	0.861	0.437
control c1 (verbal)	0.293	0.078	2.209	0.153	0.151	0.070	1.649	0.214
control c4 (verbal exec)	0.303	0.010	0.280	0.603	0.151	<0.001	<0.001	0.995

From this analysis it is clear that with prior entry of the VSR , there are no significant effects with entry of all remaining composites. However, it is possible that multiple resources are being employed in such a demanding task as COM. So one additional exploratory model was examined, with simultaneous entry of several composites.

Table 6-19. COM (groups separated; control derived composites simultaneous entry, with initial VSR entry)

	Controls				Patients			
	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
Model 1								
VSR	0.106	0.106	2.723	0.113	0.006	0.006	0.141	0.711
control c1|c4 (verbal & verbal exec)	0.272	0.166	2.391	0.116	0.117	0.110	1.312	0.290
Model 2								
VSR	0.106	0.106	2.723	0.113	0.006	0.006	2.723	0.113
control c2|c3 (both VS)	0.215	0.109	1.458	0.255	0.081	0.075	1.458	0.255
control c1|c4 (verbal & verbal exec)	0.303	0.088	1.205	0.322	0.151	0.070	1.205	0.322

As can be seen, the simultaneous entry of these variables was not statistically significant.

6.3.3.2.5 Comparison of phonological with verbal learning

One final exploratory analysis was performed to examine the verbal contribution to the three measures of the ORT. Specifically, hierarchical regression was used to compare the variance added by a phonological loop measure (digit span forward) with a verbal learning measure (Rey-AVLT total). These two measures were added following the VSR control task.

As can be seen from the analysis below, in patients, it was only the verbal learning measure which added significant variance to the OLB and COM models (models 3 to 6), irrespective of the order of entry. In controls, entry of neither measure was significant. This analysis was also repeated replacing the Rey-AVLT total with the A1 immediate recall measure and no entry step was significant (all R^2 change values <5.5%).

Finally, one additional comparison was made in patients to compare the entry of Rey-AVLT total with the delayed recall measure (A7). This was included to specifically examine the variance explained be the 'learning' aspect of the measure compared to the delayed recall component. For both OLB and COM, initial entry of the 'total' measure added significant variance while the subsequent entry of A7 did not (<3%). With initial entry of the A7 measure, entry was not statistically significant however subsequent entry of 'total' explained additional variance (OLB: 6.1% and COM: 17.7%).

Table 6-20. Comparison of a the Rey-AVLT total with digit span forward (all ORT measures)

	Controls				Patients			
	R^2	R^2 change	F for R^2 change	p	R^2	R^2 change	F for R^2 change	p
POM								
Model 1								
VSR	0.231	**0.231**	6.908	**0.015**	0.079	0.079	1.961	0.175
Rey-AVLT total	0.255	0.024	0.694	0.414	0.106	0.027	0.676	0.420
Digit span forward	0.257	0.002	0.069	0.795	0.125	0.019	0.465	0.503
Model 2								
VSR	0.231	**0.231**	6.908	**0.015**	0.079	0.079	1.961	0.175
Digit span forward	0.231	<0.001	0.006	0.938	0.099	0.021	0.502	0.486
Rey-AVLT total	0.257	0.026	0.728	0.403	0.125	0.026	0.631	0.436
OLB								
Model 3								
VSR	0.159	**0.159**	4.352	**0.048**	0.022	0.022	0.528	0.475
Rey-AVLT total	0.159	<0.001	0.010	0.922	0.262	**0.240**	7.148	**0.014**
Digit span forward	0.179	0.019	0.498	0.488	0.269	0.007	0.207	0.654
Model 4								
VSR	0.159	**0.159**	4.352	**0.048**	0.022	0.022	0.528	0.475
Digit span forward	0.178	0.018	0.494	0.490	0.032	0.009	0.215	0.648
Rey-AVLT total	0.179	0.001	0.035	0.852	0.269	**0.237**	6.825	**0.016**
COM								
Model 5								
VSR	0.106	0.106	2.723	0.113	0.006	0.006	0.141	0.711
Rey-AVLT total	0.203	0.097	2.678	0.116	0.206	**0.200**	5.527	**0.028**
Digit span forward	0.203	<0.001	0.006	0.940	0.258	0.052	1.478	0.238
Model 6								
VSR	0.106	0.106	2.723	0.113	0.006	0.006	0.141	0.711
Digit span forward	0.118	0.012	0.296	0.592	0.064	0.058	1.351	0.258
Rey-AVLT total	0.203	0.085	2.250	0.148	0.258	**0.194**	5.498	**0.029**

6.4 Discussion of Chapter 6

6.4.1 Summary of findings

This final empirical chapter presents the results of a series of regression analyses examining the relationship between spatial memory measures (SWM and ORT) and the composite measures derived from the PCAs in Chapter 4. Additional HPA axis assessments are also presented and their relationship to the above measures.

The PCA composites from Chapter 4 were used in regression models to predict separately SWM between search errors (BSE) and within search errors (WSE). There were clear and interpretable differences in the set of PCA composites which predict BSE and WSE. As regards BSE, in the whole group, and also in the control and patient groups separately, the visuospatial composite and the verbal memory composite both explained significant additional variance, irrespective of the order of entry (although the visuospatial always explained numerically more). As regards WSE, in the whole group, entry of the verbal composite was significant when entered first, but not when entered following the visuospatial composite. In controls separately, exactly the same pattern emerged, while for patients, the pattern was even more clear in that the verbal composite was not significant in any model. Interestingly, the visuospatial composite was significant only when it was derived from the patients' own PCA analysis result within this WSE analysis suggesting that, for this measure at least, the use of the more tightly defined control-derived composite was not appropriate.

As discussed earlier, the ORT yields three measures: POM, OLB and COM. With regard to POM, controls and patients showed an almost identical pattern of loading for POM in each of the analyses, with the visuospatial composite explaining a significant proportion of the variance. As regards OLB, in controls it was the verbal executive measure (verbal fluency tests) which provided the significant entry in the models, while for patients it was the verbal

composite which explained additional variance wherever it was included in the model. Due to the complexity of COM, the pattern of loadings within the models was less clear.

In patients, the pattern of loadings between SWM WSE and BSE was similar to that seen between ORT POM and OLB measures, with WSE and POM having high visuospatial loading and BSE and OLB having high verbal composite loadings. However, the visuospatial composite made a significant contribution to BSE, but not to OLB.

Finally, some exploratory analyses were included to look at specific aspects of the Rey-AVLT and the relationship with the ORT measures. It seems that specifically it was the total learning measure which related to the OLB/COM processes – neither immediate or delayed recall measures (single list) from the Rey-AVLT nor forward digit span predicted significant variance. The implications of this specificity are discussed in section 7.3.2 below.

As outlined in the introduction, with regard to the neuroendocrine assessment in this chapter, the aim was to provide a more comprehensive assessment with a focus on measures which have been shown to be sensitive and linked more closely to the function of the GR than simple peripheral measures. In Chapter 2, the assessment of levels through the afternoon resulted in a significant group difference but there was little relationship with neuropsychological functions. However, in Chapter 3, using a GR antagonist there was a clear change in neuropsychological function (specifically CANTAB SWM BSE), therefore the aim here was to attempt to examine the relationship between specific spatial processes and these more selective HPA/GR measures. Although there were no differences between the groups in salivary cortisol or DHEA measures, the proportion of patients characterised as DST non-suppressors fits exactly with that described in the original work on the development of the DST: Carroll (1982a) described sensitivity of 50-65%, although specificity was higher than that

seen here. When examining the relationship between these measures and the spatial tests, there was little evidence of any strong relationships. The only measure which was significant in the whole group (CAR AUC vs COM) explained some variance in the regression model but this was not sufficient to render the subsequent entry of the group variable non-significant.

To avoid repetition, more detailed discussion of this final empirical chapter, especially with regard to the implications of these findings, will be covered in the General Discussion (see section 7.3 below). However, some methodological issues will be briefly outlined here.

6.4.2 Methodological issues

It was noted in Section 6.3.3.2.2 that the labels used to describe each of the composites was only for general descriptive purposes only. For some, for example verbal memory, the variables loading to this composite were very consistent and well defined. However others, such as the 'visuospatial immediate' or 'visuospatial complex' the boundaries defining these are less clear, the former for example contained both spatial span measures (forward and reverse), but also the DSST (a test of psychomotor speed and attention). It should also be noted that the measures within composites (e.g. visuospatial) often cover a range of processes from memory measures to those falling broadly under the executive function domain. Therefore the processes represented or measured by a composite can be broad. A further note on the loading of the variables within each composite is that while these produced orthogonal solutions, some potentially interesting tests may have been lost to the composites through multiple loading. As discussed already, this overlap of processes may reflect the norm in human neuropsychology, therefore removing such tests to achieve independent composites may be somewhat artificial.

It should also be noted that some of the analyses performed in this last chapter were on relatively small samples, especially the sub-group of patients and controls who completed the ORT and also the numbers completing the HPA axis assessments. Some moderate R^2 change values may not therefore have reached statistical significance, although it is the case that the *relative* comparison of the pattern of loadings between the measures from the ORT is equivalent. The issue with missing or incomplete neuroendocrine samples was outlined in section 6.3.1. While some values can reasonably be imputed (e.g. CAR – estimation of the midpoint between two valid samples), other sampling protocols which involved less frequent sampling are not amenable to this without artificially decreasing the variance in the sample (e.g. replacement with the group mean).

In the last chapter of the thesis, a general discussion is presented to summarise the principal findings of the thesis and to highlight the main strengths and weaknesses, as well as outline the implications and future directions for research.

Chapter VII

General Discussion

7. General discussion

In this final chapter, the results and principal findings of Chapters 1 to 6 will be summarised before discussing the major strengths and weaknesses of these and establishing directions for future research.

7.1 Summary of principal findings in the thesis

In Chapter 1, a general background to the thesis was presented along with an overview of bipolar disorder outlining the key areas of interest to the present work, namely that patients with mood disorder often exhibit neuropsychological impairment and HPA axis dysfunction. It was noted that there was relatively little data specifically examining neuropsychological function in bipolar depression. Chapters 2 and 3 reported the results from an initial sample of 20 patients with bipolar depression. In Chapter 2, a broad profile of neuropsychological impairment was observed in patients within all the general neuropsychological domains examined, with large effects in immediate memory (phonological and spatial), executive functioning (spatial working memory, phonological fluency and sustained attention), spatial (but not visual) recognition memory, delayed verbal recall and recognition, and psychomotor speed. Sampling of cortisol and DHEA levels in plasma showed that hypercortisolaemia was evident in patients, but with no significant difference in DHEA or cortisol-DHEA ratios. However, there were no meaningful significant correlations between the extent of HPA axis disturbance and neuropsychological functioning, despite the acceptable (ROC) sensitivity and specificity of one of the HPA axis measures (plasma cortisol AUC; 1pm to 4pm). In Chapter 3, these patients entered into a double-blind, randomized crossover treatment study of the GR antagonist mifepristone which was administered adjunctively to concurrent medication. At the primary endpoint for the study (two weeks after cessation of treatment) a significant

improvement in CANTAB Spatial Working Memory (SWM)[28] between search errors was observed. At this point there was also a significant reduction in cortisol levels from baseline in the active arm of the study. There was no significant difference in symptoms between treatment arms of the study at this point. Significant relationships were observed between the primary neuropsychological and HPA axis measures in that the percentage improvement in SWM correlated significantly with the cortisol area-under-the-curve at baseline. There were also selective correlations between the degree of cortisol reduction and the extent of neuropsychological improvement in some measures.

At this point, several questions had emerged from the data, especially the improvement in one aspect of spatial working memory in response to a manipulation of the HPA axis. In order to explore these findings in more detail, a larger cohort of participants was recruited and a more comprehensive analysis of neuropsychological functioning was performed. In view of the SWM result mentioned above, this focussed particularly on additional executive and visuo-spatial memory measures. In Chapter 4, a principal components analysis (PCA) was carried out to examine the component loadings of the neuropsychological measures for the whole group and for patients and controls separately. Although the primary aim of this chapter was to generate independent components from which composites could be derived for use in later analyses, two important facets became apparent in the data. Firstly, some of the individual variables (especially visuospatial ones) loaded onto more than one composite, suggesting both that multiple interdependent processes may underpin such variables; and also, while it may be possible to extract independent orthogonal solutions from neuropsychological variables, this may not be the true 'structure' underlying some tests. The implication is that results on any one of these tests cannot be attributed to a single component. Secondly, it appeared from the general profile of these composites that while in

[28] For the remainder of this discussion, SWM will be used to refer specifically to the CANTAB test. Any other reference to spatial memory refers to the general concept or process.

controls the pattern followed clear, expected distinctions (i.e. verbal memory, verbal executive, visuospatial etc.), in patients the pattern of loadings was less clear with most components containing at least one 'verbal' measure. In Chapter 5, a more detailed analysis of spatial memory processes was performed through the administration of the novel Object Relocation Task (ORT) paradigm which permits the fractionation of independent aspects of memory: exact positional information (POM), object-to-location binding (OLB) and a combined condition (COM). Through the administration of the secondary battery of tests, relationships were observed between verbal memory and both OLB and COM processes, which along with other analyses suggested that these latter ORT measures formed a component of visuospatial memory (distinct from POM) which was processed differently in patients compared to controls.

Finally, in Chapter 6, using composites derived from the PCA solutions in Chapter 4, a series of hierarchical regression models were employed to examine their relationship with the principal spatial memory measures. Also included in this section were a number of additional measures of HPA axis function and their relationship with the spatial measures was also examined. In terms of between group differences, saliva measures did not reveal differences between patients and controls although the proportion of patients characterised as DST non-suppressors was consistent with expectations. However, little evidence of a relationship between these measures and any measure from either SWM and ORT was found. With regard to the neuropsychological composites, regression analyses produced consistent results for SWM with both patients and controls relying on both visuospatial and verbal processes for BSE, but visuospatial only for WSE. For the ORT there were interesting differences between the groups. Controls relied on visuospatial measures for POM, but verbal executive measures for OLB; while patients, although similar to controls for POM, relied much more on verbal measures for OLB. Overall, the large between group effect seen in POM remained significant

after accounting for the variance explained by multiple different composites. As discussed in section 5.4.1 this, along with the similar profile of composite loadings in the separate groups, may suggest that either that patients have a core deficit in perceptual-coordinate spatial processing or they are unable to support or maintain a precise spatial representation by some other means (either verbal/categorical coding or accurate perceptual-categorical recoding).

7.2 Methodological strengths and limitations

With regard to the neuropsychological aspects of the thesis, there are several notable strengths. Many previous studies have utilised generic test batteries to examine broad neuropsychological processes in bipolar disorder. In the present work, after establishing effects using tests which examine broad functions, more specific, novel tasks (such as the first use of ORT in bipolar disorder) were used to focus on spatial processes in more detail. This was accompanied by statistical approaches that go beyond those which are usually applied in order to understand the processes underlying each of the measures of interest. The derivation of composite scores based on the outcome of principal component analysis to aid in this interpretation was also a particular strength, going beyond the treatment of tests as individual measures and instead examining more integrated processes. However, as noted previously, one aspect of this is also a potential weakness – composite scores could be derived from oblique solutions but by deriving fully orthogonal solutions in order to achieve the greatest degree of independence between composite predictors, a certain degree of artificiality may have resulted with some tests failing to be included through loadings to multiple composites. This situation (of greater dependence between components) may be more representative of the norm with regard to complex neuropsychological tasks drawing on and tapping multiple underlying processes. It should also be noted that some of the composites represent a somewhat broad, complex range of processes, especially in the example of the visuo-spatial

composites, they were entered together into the regression models. In this latter example, the measures loading into these composites represent a range of immediate memory, longer term memory and executive/attentional processes.

One additional issue to consider is that of sensitivity to change of the neuropsychological tests used. The specificity of any observed deficit needs to be considered not just in terms of the magnitude of differences (as discussed regarding the differential deficit analysis) but also in terms of the 'discriminating power' of a test as reflected in its reliability, the shape and distribution of scores and item difficulty (Chapman & Chapman, 1973; Chapman & Chapman, 1978). This same argument could be applied to sensitivity of a test (or measure) to change, although it should be noted that SWM BSE was not the largest effect size (between patients and controls) found at baseline therefore it does not appear to be the case that this measure changed because had the greatest chance of improvement. Also the crossover design of the study meant that comparisons were made within subject, including treatment conditions.

With regard to the neuroendocrine tests, general issues around sample size and missing data were highlighted at the end of Chapter 6. To focus first on these particular methods of HPA axis assessment, in Chapter 2 a multi-time point blood sampling method was adopted which was taken over the period of neuropsychological assessment. In the later studies in the thesis, other methods were utilised, including salivary 8am/8pm cortisol-DHEA ratio, salivary CAR and the DST which have been described as sensitive measures in the literature. However, it is not known why these alternative saliva methods did not reveal any between group differences despite a high proportion of patients being DST non-suppressors (of course it is noted that in the case of the latter, the group difference is unsurprising since it is based on a cut-point which is specifically derived to maximise the difference between patients and control groups).

There are examples in the literature where one type of sampling methodology has failed to find statistical differences between groups while another has, within the same individuals. For example, differences in CAR have been described in individuals with no evidence of DST non-suppression (Deshauer et al., 2003), while in other studies pronounced DST and dex/CRH abnormalities have been found in individuals with no differences in salivary cortisol levels (e.g. Watson et al., 2004). In this case, several different methodological issues may have contributed to the discrepancy. Firstly, the first sample used plasma levels over multiple afternoon time-points and the second, single point saliva levels. Secondly, the initial sample was recruited from patients in a tertiary referral service who presumably had more complex long-standing illnesses, while the latter sample also recruited directly from secondary care. Thirdly, it may have been a result of switching from the precision of collecting samples in the laboratory to participants collecting their own at home. CAR for example is affected by the actual awakening time (Federenko et al., 2004). In a recent meta-analysis which assessed twenty case-control studies where salivary cortisol was examined, including 1354 patients with depression (unipolar and bipolar) and 1052 controls, concluded that although there were very small increases in morning and evening cortisol levels in patients, overall it was a poor discriminator. In this meta-analysis, substantial overlap in levels between the groups was observed as well as high heterogeneity in morning sampling. Factors such as higher intra-assay coefficients of variation in cortisol kits and mean age were associated with a higher mean difference in morning salivary cortisol between depressed and controls, while the variables 'gender' and 'depression severity' were not (Knorr et al., 2010).

7.3 Implications of the primary findings

In the next section, a number of key implications of the present work are discussed and compared to previous work.

7.3.1 What is the relationship between the HPA axis and spatial memory?

The results of the present thesis provide little evidence of a simple relationship between peripheral HPA axis measures and spatial memory. While such measures are certainly useful to establish if neuroendocrine abnormalities are present on a group level, the variability inherent from the pulsatility of the HPA axis to the methodological issues discussed make it difficult to reliably use such indices to assess the link. However, clear effects on SWM were seen when the HPA axis was directly targeted by the GR antagonist mifepristone. This is in line with previous animal work showing effects of mifepristone on spatial memory (Oitzl et al., 1998a).

One feature of the results of Chapter 3 which is of particular note (and will be discussed in more detail below) is that while SWM BSE was improved, there was no significant change in verbal memory as assessed by the Rey-AVLT. This of course may simply be due to differences in test sensitivity to change. However it may also be due to the processes underpinning SWM BSE. One speculative explanation is that avoiding SWM BSE may be a somewhat 'categorical' process (or at least either overlaps with the same processes as categorical-type measures or becomes so as the test proceeds and participants form more accurate spatial categories relevant to the successful execution of the search sequence) in that all the spatial locations on the screen are fixed and the participant must search through these for the target, which

involves a (cognitively demanding) updating of each target location[29]. This explains the high visuospatial composite loading but also the verbal composite contribution (section 6.3.3.1). As has been discussed, the 'verbal' loading in this context may not be attributable to linguistic properties of the test, but to a categorical aspect. Therefore it may be that categorical spatial processes are sensitive to the effects of the GR antagonist i.e. SWM BSE were affected but not WSE. And the lack of effect of the drug on the Rey-AVLT may support the contention that it is not the verbal/linguistic elements of the task which is related to the BSE but in fact the categorical facets underpinning it. This may be similar for ORT OLB/COM measures (see section 7.4 below).

7.3.2 Verbal scaffolding of visuospatial measures in bipolar depression?

There have been a small number of studies that have examined the effects of verbal mediation on object-location memory processes, including some utilising similar paradigms to the present study (see Sections 5.1 and 5.4). In his seminal paper, Kosslyn suggests a *relative* hemispheric specialisation for categorical and co-ordinate spatial relations. Following initial experiments to establish the basic principle of a left hemisphere advantage for categorical and right hemisphere advantage for co-ordinate processes, explores the effects of repetition, difficulty and categorical dimension (i.e. on/off, above/below) on performance. With repeated presentation, the initial right hemisphere advantage reduced. It was argued that this was because practice allowed the development of new finer grained categories able to capture these spatial relations. Once these new categories were formed it was possible for the metric judgment task to be resolved using categorical processing and the encoding of spatial information with respect to fine grained metric distance would be diminished. However, a corresponding left hemisphere advantage did not develop. It was argued, therefore, that that

[29] Of course the temporal aspect of this type of error monitoring is also noted i.e. to avoid BSE, target locations are retained over longer periods of time than required to avoid WSE.

this may not be a direct verbal/semantic effect *per se*, but possibly a disruption in underlying categorical information processing (Kosslyn et al., 1989).

The earlier discussion in Chapter 5 outlines comparable thinking with regard to the mediation of the ORT OLB and COM measures i.e. that the mediation may not be specifically linguistic. The original Postma and de Haan (1996) paper reported that articulatory suppression (AS+) had effects on tasks using stimuli that are not readily nameable and suggested that this may not be due to disruption of verbal processing *per se*, but possibly a disruption in underlying categorical information processing. One important feature of the studies examining the effect of verbal mediation of processes within the ORT is that they all used AS+ (either counting or repetition of a nonsense syllable) as the concurrent task. The aim of this is to interfere with the functioning of the phonological loop, without placing demands on executive processes (which can occur if the task is too complex). In the present thesis, one exploratory analysis was added to section 6.3.3.2.5 to compare the variance explained by a 'phonological loop (PL) measure' (digit span) compared to verbal learning. In patients, while neither measure was significant in the POM analysis, both OLB and COM measures had significant variance explained by the entry of verbal learning[30], but not the PL measure. Although these two methods (i.e. comparing the effects of a direct interference task with the relationship in regression models) are of course very different, together they may further strengthen the argument that it may not be the verbal/linguistic element, but in fact the learning or categorical coding of information that is key. Of course, although the previous work described suggests that it is not the linguistic element *per se* that is responsible for the results observed, it cannot be completely ruled out that it is a simple verbal coding method that is being attempted by the patients. One final point to note is that the evidence for AS+ effects are not

[30] n.b. this analysis was also repeated, replacing Rey-AVLT total with immediate verbal recall (trial A1 of the Rey-AVLT) or delayed recall (recall of list A7) and no significant entry steps were evident for either measure suggesting that this is not the mediating factor.

always consistent in terms of which ORT processes are affected (see section 5.1), therefore this argument remains speculative in the context of the present data.

Considering the above discussion in the context of Kemmerer and Tranel's (2000) work on the dissociation between linguistic and perceptual representations of categorical spatial relationships, it could be hypothesised that there is a dissociation between patients and healthy controls in the relative importance of the two, with patients either having greater dependence on the linguistic (verbal-categorical) or else using this to scaffold performance. If this is the case, then the explanation for the pattern of impairment seen in the patients in the ORT (i.e. arithmetically larger effects for POM over OLB/COM and the removal of the group effect in the latter when factoring in verbal measures, as well as the pattern of regression loadings) is that all measures essentially involve encoding and maintaining a complex spatial representation. Achieving this through linguistic representation is obviously less precise than through perceptual processes, therefore whilst the former may to some degree aid OLB performance – and to a lesser extent – COM, it cannot match the precision of the latter. In the case of POM where all items are identical, linguistic processes cannot aid performance in the same way. It may be that patients do not use verbal processes at all for precise co-ordinate tasks, or that they use them, but to no effect. In either case verbal processes will not relate to or explain significant variance in POM task performance.

7.3.3 Stability of co-ordinate representations – an inefficiency of categorical coding in bipolar patients?

The POM measure (from the ORT) yielded large effect sizes between patients and controls. For both groups, this measure had a strong association with the visuospatial composites (see section 6.3.3.2.2). In multiple regression models to predict POM, 'group' continued to be a

significant predictor of POM even after PCA-derived composites had been entered. As discussed previously, the maintenance of precise, fine-grain representation, as assessed by POM, stands apart from other spatial location measures. Such representation appears to be highly sensitive to retention duration and exhibits a rapid decay profile. Several authors have argued that the fragility of these representations require a rapid re-coding into categorical representation to facilitate maintenance over time (see discussion in Section 5.4, p.256). It is possible that such a recoding process is impaired in BD and leads to the large impairment in POM observed in these bipolar patients relative to controls.

It is worth noting that the level of performance in patients does not approach anywhere near floor effects (i.e. the level of accuracy determined as chance- Kessels et al., 2000b; Postma et al., 2000). Therefore, as discussed at the end of Section 5.4, this deficit may represent a failure to rapidly process and integrate into alternative categorical coding. It would be expected that such processes would have significant executive contributions and as was observed in the composite loadings, both visuospatial composites would together predict significant variance. These composites included visuospatial executive measures such as SWM strategy (the efficiency of deriving and employing a spatial search strategy), SOPT (the ability to generate and monitor a sequence of responses i.e. executive control of WM), reverse spatial span (executive control of visuospatial WM) and DSST (rapid visuospatial processing/ psychomotor speed). Interestingly in the hierarchical regression models, the relationship of the composites to POM was similar in patients and controls, perhaps leading to the size of the between group difference, as alternative categorical processes could not be relied upon as they were in OLB/COM.

7.4 Future directions and further research

Several important and interesting areas for future research have been identified, based on the present work. Some of these findings can lead straight to direct, testable hypotheses.

One area relates to the effects of GR antagonists on spatial memory. In Chapter 3, results indicated that at the point where the improvement in spatial working memory was observed there was no significant difference in mood therefore minimizing this potential confound to interpretation. In the study, the neuropsychological assessment was performed 2 weeks after cessation of treatment to avoid the acute effects of elevated cortisol levels. However, it is important to understand the full timeframe of the effects of the GR antagonist. For example, there may be multiple mechanisms at work leading to the effect – from the animal work discussed in Section 1.5, GR antagonists are seen to have a direct effect on spatial memory, therefore at this point it is unknown if the SWM improvement was the result of the drug effect at the GR (which due to its long half-life would still be present) or an alteration in HPA axis function and lowering of cortisol levels. One final alternative explanation is based on the notion many antidepressant drugs have actions on blood-brain barrier steroid transporters (such as multidrug resistance p-glycoprotein). Plasma cortisol cannot freely enter the brain by passive diffusion because its access is limited by such membrane steroid transporters which actively expel cortisol from the brain. It has been suggested that by inhibiting membrane steroid transporters at the blood–brain barrier and in neurones, more cortisol is able to enter the brain (Pariante, 2004; Pariante et al., 2004b), thereby restoring glucocorticoid-mediated negative feedback of the HPA axis (Pariante et al., 2004b). Hypercortisolaemia is therefore argued to be a possible compensatory adaptive response to a central *hypo*cortisolemic state (Pariante, 2003). Considering mifepristone: the antagonist action of mifepristone on GR causes a robust (2- to 3-fold) elevation in cortisol levels and this may facilitate HPA axis negative feedback.

As highlighted in Chapter 4, the patients who were recruited for this latter section of the thesis (Chapters 4 to 6) were part of a program of research further examining the effects of GR antagonists in bipolar disorder. This research used the same drug protocol as in Chapter 3, but a more simple parallel group design was used. As part of this, all neuropsychological tests, including the ORT paradigm, were administered at the point of cessation of treatment as well as 2 weeks later. Therefore there is the opportunity to study the timeframe of the effects on memory. Specific hypotheses can be made regarding the role of cortisol, for example, the high cortisol levels at the point of cessation of treatment would be expected to impair performance. However, if it is a direct drug effect which is important, improvements could be expected immediately. Further interesting hypotheses can be made regarding the specific processes which will be affected. For instance, as it was the SWM BSE which was improved (this is the measure with verbal loadings, in contrast to WSE) it can be hypothesised that OLB and COM measures of the ORT should be similarly improved. It would also be of interest to further examine these effects in euthymic patients as well as in healthy subjects, to see what effect GR manipulation produced on tasks where identical patterns of composite loadings were observed between patients and controls (e.g. SWM and POM) compared with those where differences were observed (e.g. OLB).

As a next step, understanding the precise brain structures involved in these processes following HPA axis manipulation is important. Throughout this thesis, the role of structures like the PFC, hippocampus and parietal lobes have been mentioned with respect to both the effects of cortisol and structures that mediate performance on the spatial memory processes of interest. The effects of the GR antagonists are particularly related to this – as discussed, the animal work has shown mifepristone to have effects on hippocampus as well as prefrontal cortex corticosteroid receptor levels (Bachmann et al., 2003). Returning to the earlier discussion of the brain structures underpinning different spatial memory processes (e.g.

Section 3.4.1 and 5.1.3) it may be hypothesised that specific aspects of some tests will be affected by GR manipulation, for example, we could expect that SWM between search errors and ORT OLB (perhaps COM) would be affected if actions at the level of the hippocampus are important. Imaging and electroencephalogram studies may well prove invaluable in this respect. A number of studies using these techniques to explore categorical and co-ordinate processes within WM have been conducted (e.g. van der Lubbe et al., 2006; van der Ham et al., 2009) and these may be of interest if applied to bipolar disorder or following HPA axis manipulation.

With regard to future experimental neuropsychology studies, exploring the effects of verbal interference on the SWM and ORT measures in bipolar patients is of great interest. While some studies have shown effects of AS+ on OLB/COM processes but not POM (e.g. Postma & de Haan, 1996; Kessels & Postma, 2002), results of other studies have suggested that unexpected inconsistencies can occur with concurrent AS+ (Postma et al., 1998; Postma et al., 1999), possibly as a result of interaction with other methodological factors (see section 5.1.1). Therefore, either more work is needed to first optimise the AS+ task to avoid placing demands on executive resources, or the use of such tasks are inherently problematic due to the reasons discussed earlier on the relative contribution of linguistic and perceptual representations of visuospatial arrays and the executive demands of exact, metric detail (see sections 7.3.2 and 7.3.3 above). Exploring alternative ways of interfering with different elements of the linguistic/perceptual representations may therefore offer a novel method of examining this phenomenon.

A final area which should be further explored is the temporal aspect of visuospatial representation, specifically with the ORT task. As has been discussed, questions remain such as whether it is the case that because exact coordinate representation is so resource-

demanding it is quickly reduced to a categorical representation? Is it the case that controls have a better ability to recode the coordinate to categorical and therefore do better or have more accurate categorical representation? Adopting methodologies such as those used by McNamara et al. (1992) or Werner & Diedrichsen (2002) to examine memory for precise location and temporal decay, and comparing patients and controls, would prove interesting. One related planned analysis using the data from the present study is to examine time vs. accuracy plots for the ORT paradigm data to examine the decay curve in the healthy controls (to ascertain if a limited number of items are very accurately located then the remainder of the array is more grossly arranged around these reference points) and then compare these results to the data from the bipolar group.

Carrying out these additional studies will further our knowledge of the integration among the biological mechanisms underlying neuropsychological impairment in mood disorders and should develop our understanding of integration between cognitive processes in general.

References

8. References

Aas, M., Dazzan, P., Mondelli, V., Toulopoulou, T., Reichenberg, A., Di Forti, M., Fisher, H. L., Handley, R., Hepgul, N., Marques, T., Miorelli, A., Taylor, H., Russo, M., Wiffen, B., Papadopoulos, A., Aitchison, K. J., Morgan, C., Murray, R. M., & Pariante, C. M. (2010). Abnormal cortisol awakening response predicts worse cognitive function in patients with first-episode psychosis. *Psychological Medicine, FirstView*, 1-14.

Adler, G., & Jajcevic, A. (2001). Post-dexamethasone cortisol level and memory performance in elderly depressed patients. *Neuroscience Letters, 298*(2), 142-144.

Akiskal, H. S., & Pinto, O. (1999). The evolving bipolar spectrum. Prototypes I, II, III, and IV. *Psychiatr Clin North Am, 22*(3), 517-534, vii.

Alessio, A., Bonilha, L., Rorden, C., Kobayashi, E., Min, L. L., Damasceno, B. P., & Cendes, F. (2006). Memory and language impairments and their relationships to hippocampal and perirhinal cortex damage in patients with medial temporal lobe epilepsy. *Epilepsy & Behavior, 8*(3), 593-600.

Altman, D. G., Machin, D., Bryant, T. N., & Gardner, M. J. (2000). *Statistics with confidence* (2nd ed.). Bristol: BMJ books.

Anacker, C., Zunszain, P. A., Carvalho, L. A., & Pariante, C. M. The glucocorticoid receptor: Pivot of depression and of antidepressant treatment? *Psychoneuroendocrinology, In Press, Corrected Proof*.

Angst, J., & Marneros, A. (2001). Bipolarity from ancient to modern times: conception, birth and rebirth. *Journal of Affective Disorders, 67*(1-3), 3-19.

APA. (1994). *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV* (fourth ed ed.). Washington, DC.: American Psychiatric Association.

Appelhof, B. C., Huyser, J., Verweij, M., Brouwer, J. P., van Dyck, R., Fliers, E., Hoogendijk, W. J. G., Tijssen, J. G. P., Wiersinga, W. M., & Schene, A. H. (2006). Glucocorticoids and

relapse of major depression (dexamethasone/Corticotropin-Releasing Hormone test in relation to relapse of major depression). *Biological Psychiatry, 59*(8), 696-701.

Ashok, P. W., Wagaarachchi, P. T., & Templeton, A. (2002). The antiprogestogen mifepristone: a review. *Current Medicinal Chemistry - Immunology, Endocrine & Metabolic Agents, 2*(2), 71-90.

Assies, J., Visser, I., Nicolson, N. A., Eggelte, T. A., Wekking, E. M., Huyser, J., Lieverse, R., & Schene, A. H. (2004). Elevated salivary dehydroepiandrosterone-sulfate but normal cortisol levels in medicated depressed patients: preliminary findings. *Psychiatry Research, 128*(2), 117-122.

Aubry, J.-M., Gervasoni, N., Osiek, C., Perret, G., Rossier, M. F., Bertschy, G., & Bondolfi, G. (2007). The DEX/CRH neuroendocrine test and the prediction of depressive relapse in remitted depressed outpatients. *Journal of Psychiatric Research, 41*(3-4), 290-294.

Bachmann, C. G., Linthorst, A. C., Holsboer, F., & Reul, J. M. (2003). Effect of chronic administration of selective glucocorticoid receptor antagonists on the rat hypothalamic-pituitary-adrenocortical axis. *Neuropsychopharmacology, 28*(6), 1056-1067.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417-423.

Baddeley, A., & Wilson, B. (1988). Frontal amnesia and the dysexecutive syndrome. *Brain and Cognition, 7*(2), 212-230.

Baddeley, A. D., Emslie, H., & Nimmo-Smith, I. (1992). *The Speed and Capacity of Language Processing (SCOLP) test*. Bury St Edmunds, Suffolk: Thames Valley Test Company.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8). New York: Academic Press.

Basso, M. R., Lowery, N., Neel, J., Purdie, R., & Bornstein, R. A. (2002). Neuropsychological impairment among manic, depressed, and mixed-episode inpatients with bipolar disorder. *Neuropsychology, 16*(1), 84-91.

Baulieu, E.-E., & Robel, P. (1998). Dehydroepiandrosterone (DHEA) and dehydroepiandrosterone sulfate (DHEAS) as neuroactive neurosteroids. *PNAS, 95*(8), 4089-4091.

Baulieu, E. E. (1981). Steroid hormones in the brain: Several mechanisms? In K. Fuxe & J. A. Gutafsson & L. Wetterberg (Eds.), *Steroid Hormone Regulation of the Brain* (pp. 3-14). Oxford: Pergamon Press.

Baulieu, E. E. (1997). Neurosteroids: of the nervous system, by the nervous system, for the nervous system. *Recent Prog Horm Res, 52*, 1-32.

Baulieu, E. E., & Robel, P. (1996). Dehydroepiandrosterone and dehydroepiandrosterone sulfate as neuroactive neurosteroids. *J Endocrinol, 150 Suppl*, S221-239.

Bearden, C. E., Hoffman, K. M., & Cannon, T. D. (2001). The neuropsychology and neuroanatomy of bipolar affective disorder: a critical review. *Bipolar disorders, 3*(3), 106-150.

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.

Belanoff, J. K., Rothschild, A. J., Cassidy, F., DeBattista, C., Baulieu, E.-E., Schold, C., & Schatzberg, A. F. (2002). An open label trial of C-1073 (mifepristone) for psychotic major depression. *Biological Psychiatry, 52*(5), 386-392.

Berne, R. M., & Levy, M. N. (1998). *Physiology* (Fourth ed.). St. Louis: Mosby Inc.

Bhagwagar, Z., Hafizi, S., & Cowen, P. J. (2003). Increase in concentration of waking salivary cortisol in recovered patients with depression. *Am J Psychiatry, 160*(10), 1890-1891.

Bhagwagar, Z., Hafizi, S., & Cowen, P. J. (2005). Increased salivary cortisol after waking in depression. *Psychopharmacology, 182*(1), 54-57.

Bjornerem, A., Straume, B., Oian, P., & Berntsen, G. K. R. (2006). Seasonal variation of estradiol, follicle stimulating hormone, and dehydroepiandrosterone sulfate in women and men. *J Clin Endocrinol Metab, 91*(10), 3798-3802.

Bland, J. M., & Altman, D. G. (1996). Statistics Notes: Transforming data. *BMJ, 312*(7033), 770-.

Board, F., Persky, H., & Hamburg, D. A. (1956). Psychological stress and endocrine functions; blood levels of adrenocortical and thyroid hormones in acutely disturbed patients. *Psychosomatic Medicine, 18*, 324-333.

Board, F., Wadeson, R., & Persky, H. (1957). Depressive affect and endocrine functions; blood levels of adrenal cortex and thyroid hormones in patients suffering from depressive reactions. *AMA Arch Neurol Psychiatry, 78*(6), 612-620.

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*(1-2), 1-20.

Borkowska, A., & Rybakowski, J. K. (2001). Neuropsychological frontal lobe tests indicate that bipolar depressed patients are more impaired than unipolar. *Bipolar Disorders, 3*(2), 88-94.

Bourgeois, S., Pfahl, M., & Baulieu, E. E. (1984). DNA binding properties of glucocorticosteroid receptors bound to the steroid antagonist RU-486. *EMBO J., 3*(4), 751-755.

Brambilla, D. J., O'Donnell, A. B., Matsumoto, A. M., & McKinlay, J. B. (2007). Lack of seasonal variation in serum sex hormone levels in middle-aged to older men in the Boston area. *J Clin Endocrinol Metab, 92*(11), 4224-4229.

Bremner, J. D., Vythilingam, M., Vermetten, E., Anderson, G., Newcomer, J. W., & Charney, D. S. (2004). Effects of glucocorticoids on declarative memory function in major depression. *Biological Psychiatry, 55*(8), 811-815.

Brooks, J. O., Wang, P. W., Strong, C., Sachs, N., Hoblyn, J. C., Fenn, R., & Ketter, T. A. (2006). Preliminary evidence of differential relations between prefrontal cortex metabolism

and sustained attention in depressed adults with bipolar disorder and healthy controls. *Bipolar Disorders, 8*(3), 248-254.

Browne, E. S., Wright, B. E., Porter, J. R., & Svec, F. (1992). Dehydroepiandrosterone: antiglucocorticoid action in mice. *Am J Med Sci, 303*, 366-371.

Burdick, K. E., Gunawardane, N., Goldberg, J. F., Halperin, J. M., Garno, J. L., & Malhotra, A. K. (2009). Attention and psychomotor functioning in bipolar depression. *Psychiatry Research, 166*(2-3), 192-200.

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience, 12*(1), 1-47.

Cahill, L. (2006). Why sex matters for neuroscience. *Nature Reviews Neuroscience, 7*(6), 477-484.

Calev, A., Korin, Y., Shapira, B., Kugelmass, S., & Lerer, B. (1986). Verbal and non-verbal recall by depressed and euthymic affective patients. *Psychol Med, 16*(4), 789-794.

Callies, F., Arlt, W., Siekmann, L., Hubler, D., Bidlingmaier, F., & Allolio, B. (2000). Influence of oral dehydroepiandrosterone (DHEA) on urinary steroid metabolites in males and females. *Steroids, 65*(2), 98-102.

Campolongo, P., Roozendaal, B., Trezza, V., Hauer, D., Schelling, G., McGaugh, J. L., & Cuomo, V. (2009). Endocannabinoids in the rat basolateral amygdala enhance memory consolidation and enable glucocorticoid modulation of memory. *PNAS, 106*(12), 4888-4893.

Cardounel, A., Regelson, W., & Kalimi, M. (1999). Dehydroepiandrosterone protects hippocampal neurons against neurotoxin-induced cell death: mechanism of action. *Proc Soc Exp Biol Med, 222*(2), 145-149.

Carroll, B. J. (1982a). Clinical applications of the dexamethasone suppression test for endogenous depression. *Pharmacopsychiatria, 15*(1), 19-25.

Carroll, B. J. (1982b). The dexamethasone suppression test for melancholia. *Br J Psychiatry, 140*(March), 292-304.

Carroll, B. J., Curtis, G. C., Davies, B. M., Mendels, J., & Sugerman, A. A. (1976a). Urinary free cortisol excretion in depression. *Psychol Med, 6*(1), 43-50.

Carroll, B. J., Curtis, G. C., & Mendels, J. (1976b). Cerebrospinal fluid and plasma free cortisol concentrations in depression. *Psychol Med., 6*(2), 235-244.

Carroll, B. J., Feinberg, M., & Greden, J. F. (1981). A specific laboratory test for the diagnosis of melancholia. Standardization, validation, and clinical utility. *Archives of General Psychiatry, 38*(1), 15-22.

Carroll, B. J., & Rubin, R. T. (2006). Is mifepristone useful in psychotic depression? *Neuropsychopharmacology, 31*(12), 2793-2794.

Carroll, B. J., & Rubin, R. T. (2008). Mifepristone in psychotic depression? *Biological Psychiatry, 63*(1), e1-e1.

Cegalis, J., & Bowlin, J. (1991). Vigil: software for the assessment of attention. Nashua, NH: Forthought.

Chapman, L. J., & Chapman, J. P. (1973). Problems in the measurement of cognitive deficit. *Psychol Bull., 79*(6), 380-385.

Chapman, L. J., & Chapman, J. P. (1978). The measurement of differential deficit. *Journal of Psychiatric Research, 14*, 303-311.

Chu, J. W., Matthias, D. F., Belanoff, J., Schatzberg, A., Hoffman, A. R., & Feldman, D. (2001). Successful long-term treatment of refractory Cushing's disease with high-dose mifepristone (RU 486). *Journal of Clinical Endocrinology and Metabolism, 86*(8), 3568-3573.

Clark, L., Iversen, S. D., & Goodwin, G. M. (2002). Sustained attention deficit in bipolar disorder. *British Journal of Psychiatry, 180*, 313-319.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, S. I. (1980). Cushing's syndrome: a psychiatric study of 29 patients. *Br J Psychiatry, 136*(Feb), 120-124.

Corpechot, C., Robel, P., Axelson, M., Sjovall, J., & Baulieu, E. E. (1981). Characterization and measurement of dehydroepiandrosterone sulfate in rat brain. *Proc. Natl. Acad. Sci. USA., 78*, 4704-4707.

Crawford, J. R., Blackmore, L. M., Lamb, A. E., & Simpson, S. A. (2000). Is there a differential deficit in fronto-executive functioning in Huntington's Disease? *Clinical Neuropsychological Assessment, 1*, 4-20.

Crawford, J. R., Wright, R. L., & Bate, A. (1995). Verbal, figural and ideatonal fluency in CHI (Abstract). *Journal of the International Neuropsychological Society, 1*, 321.

Curtin, F., & Schulz, P. (1998). Multiple correlations and Bonferroni's correction. *Biol Psychiatry, 44*(8), 775-777.

Czobor, P., Jaeger, J., Berns, S. M., Gonzalez, C., & Loftus, S. (2007). Neuropsychological symptom dimensions in bipolar disorder and schizophrenia. *Bipolar Disorders, 9*, 71-92.

Dagher, A., Owen, A. M., Boecker, H., & Brooks, D. J. (2001). The role of the striatum and hippocampus in planning: a PET activation study in Parkinson's disease. *Brain, 124*(Pt 5), 1020-1032.

Davidson, P. S. R., Gao, F. Q., Mason, W. P., Winocur, G., & Anderson, N. D. (2008). Verbal fluency, Trail Making, and Wisconsin Card Sorting Test performance following right frontal lobe tumor resection. *Journal of Clinical and Experimental Neuropsychology, 30*(1), 18 - 32.

de Kloet, E. R., Oitzl, M. S., & Joels, M. (1999). Stress and cognition: are corticosteroids good or bad guys? *Trends in Neurosciences, 22*(10), 422-426.

de Kloet, E. R., & Reul, J. M. H. M. (1987). Feedback action and tonic influence of corticosteroids on brain function: A concept arising from the heterogeneity of brain receptor systems. *Psychoneuroendocrinology, 12*(2), 83-105.

de Quervain, D. J., Henke, K., Aerni, A., Treyer, V., McGaugh, J. L., Berthold, T., Nitsch, R. M., Buck, A., Roozendaal, B., & Hock, C. (2003). Glucocorticoid-induced impairment of declarative memory retrieval is associated with reduced blood flow in the medial temporal lobe. *Eur J Neurosci, 17*(6), 1296-1302.

de Quervain, D. J., Roozendaal, B., Nitsch, R., McGaugh, J., & Hock, C. (2000). Acute cortisone administration impairs retrieval of long-term declarative memory in humans. *Nature Neuroscience, 3*(4), 313-314.

Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: a tool for unwelding visuo-spatial memory. *Neuropsychologia, 37*(10), 1189-1199.

Della Sala, S., Gray, C., Baddeley, A., & Wilson, L. (1997). *The Visual Patterns Test: a new test of short-term visual recall* Feltham, Suffolk: Thames Valley Test Company.

Dent, K. (2009). Coding categorical and coordinate spatial relations in visual-spatial short-term memory. *The Quarterly Journal of Experimental Psychology, 62*(12), 2372 - 2387.

Deptula, D., Manevitz, A., & Yozawitz, A. (1991). Asymmetry of recall in depression. *J Clin Exp Neuropsychol., 13*(6), 854-870.

Deshauer, D., Duffy, A., Alda, M., Grof, E., Albuquerque, J., & Grof, P. (2003). The cortisol awakening response in bipolar illness: a pilot study. *Can J Psychiatry, 48*(7), 462-466.

Diebold, K., Kick, H., & Schmidt, G. (1981). Urinary free cortisol excretion in endogenously depressed and schizophrenic patients. *Psychiatr Clin (Basel), 14*(1), 43-48.

Dixon, T., Kravariti, E., Frith, C., Murray, R. M., & McGuire, P. K. (2004). Effect of symptoms on executive function in bipolar illness. *Psychol Med, 34*(5), 811-821.

Douma, B. R. K., Korte, S. M., Buwalda, B., la Fleur, S. E., Bohus, B., & Luiten, P. G. M. (1998). Repeated blockade of mineralocorticoid receptors, but not of glucocorticoid receptors impairs food rewarded spatial learning. *Psychoneuroendocrinology, 23*(1), 33-44.

Du, J., Wang, Y., Hunter, R., Wei, Y., Blumenthal, R., Falke, C., Khairova, R., Zhou, R., Yuan, P., Machado-Vieira, R., McEwen, B. S., & Manji, H. K. (2009). Dynamic regulation of mitochondrial function by glucocorticoids. *Proceedings of the National Academy of Sciences, 106*(9), 3543-3548.

Duval, F., Mokrani, M. C., Crocq, M. A., Bailey, P. E., Diep, T. S., Correa, H., & Macher, J. P. (2000). Dopaminergic function and the cortisol response to dexamethasone in psychotic depression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry, 24*(2), 207-225.

Eacott, M. J., & Gaffan, D. (1991). The role of monkey inferior parietal cortex in visual discrimination of identity and orientation of shapes. *Behav Brain Res, 46*(1), 95-98.

Egeland, J., Lund, A., Landro, N. I., Rund, B. R., Sundet, K., Asbjornsen, A., Mjellem, N., Roness, A., & Stordal, K. I. (2005). Cortisol level predicts executive and memory function in depression, symptom level predicts psychomotor speed. *Acta Psychiatrica Scandinavica, 112*(6), 434-441.

El-Badri, S. M., Ashton, C. H., Moore, P. B., Marsh, V. R., & Ferrier, I. N. (2001). Electrophysiological and cognitive function in young euthymic patients with bipolar affective disorder. *Bipolar Disord, 3*(2), 79-87.

Epperson, C. N., Wisner, K. L., & Yamamoto, B. (1999). Gonadal steroids in the treatment of mood disorders. *Psychosom Med, 61*(5), 676-697.

Eser, D., Schule, C., Baghai, T. C., Romeo, E., & Rupprecht, R. (2006). Neuroactive steroids in depression and anxiety disorders: clinical studies. *Neuroendocrinology, 84*(4), 244-254.

Fabian, T. J., Dew, M. A., Pollock, B. G., Reynolds, C. F., Mulsant, B. H., Butters, M. A., Zmuda, M. D., Linares, A. M., Trottini, M., & Kroboth, P. D. (2001). Endogenous concentrations

of DHEA and DHEA-S Decrease with remission of depression in older adults. *Biological Psychiatry, 50*(10), 767-774.

Fairchild, G., Leitch, M. M., & Ingram, C. D. (2003). Acute and chronic effects of corticosterone on 5-HT1A receptor-mediated autoinhibition in the rat dorsal raphe nucleus. *Neuropharmacology, 45*(7), 925-934.

Falret, J. P. (1851). De la folie circulaire ou forme de maladie mentale caracterisee par l'alternative reguliere de la manie de la melancolie. *Bull. Acad. Natl. Med. (Paris)*.

Faravelli, C., Guerrini Degl'Innocenti, B., Aiazzi, L., Incerpi, G., & Pallanti, S. (1990). Epidemiology of mood disorders: a community survey in Florence. *J Affect Disord., 20*(2), 135-141.

Federenko, I., Wust, S., Hellhammer, D. H., Dechoux, R., Kumsta, R., & Kirschbaum, C. (2004). Free cortisol awakening responses are influenced by awakening time. *Psychoneuroendocrinology, 29*(2), 174-184.

Feigenbaum, J. D., Morris, R. G., & Polkey, C. E. (1996). Deficits in spatial working memory after unilateral temporal lobectomy in man. *Neuropsychologia, 34*(3), 163-176.

Feldman, S., Conforti, N., & Weidenfeld, J. (1995). Limbic pathways and hypothalamic neurotransmitters mediating adrenocortical responses to neural stimuli. *Neuroscience & Biobehavioral Reviews, 19*(2), 235-240.

Ferrari, E., Casarotti, D., Muzzoni, B., Albertelli, N., Cravello, L., Fioravanti, M., Solerte, S. B., & Magri, F. (2001a). Age-related changes of the adrenal secretory pattern: possible role in pathological brain aging. *Brain Res Brain Res Rev, 37*(1-3), 294-300.

Ferrari, E., Cravello, L., Muzzoni, B., Casarotti, D., Paltro, M., Solerte, S. B., Fioravanti, M., Cuzzoni, G., Pontiggia, B., & Magri, F. (2001b). Age-related changes of the hypothalamic-pituitary-adrenal axis: pathophysiological correlates. *Eur J Endocrinol, 144*(4), 319-329.

Ferrari, E., Mirani, M., Barili, L., Falvo, F., Solerte, S. B., Cravello, L., Pini, L., & Magri, F. (2004). Cognitive and affective disorders in the elderly: a neuroendocrine study. *Archives of Gerontology & Geriatrics - Supplement*(9), 171-182.

Ferrier, I. N., Stanton, B. R., Kelly, T. P., & Scott, J. (1999). Neuropsychological function in euthymic patients with bipolar disorder. *Br J Psychiatry, 175*, 246-251.

Field, A. (2000). *Discovering Statistics using SPSS for Windows*. London: Sage Publications.

Field, A. (2009). *Discovering Statistics using SPSS* (3rd edition ed.). London: Sage Publications Ltd.

First, M. B., Spitzer, R. L., Williams, J. B. W., & Gibbon, M. (1995). Structured Clinical Interview for DSM-IV (SCID-I), Research Version. New York: Biometrics Research Department, New York State Psychiatric Institute.

Foldesi, I., Falkay, G., & Kovacs, L. (1996). Determination of RU486 (mifepristone) in blood by radioreceptorassay; A pharmacokinetic study. *Contraception, 54*(1), 27-32.

Follenius, M., Simon, C., Brandenberger, G., & Lenzi, P. (1987). Ultradian plasma corticotropin and cortisol rhythms: time-series analyses. *J Endocrinol Invest, 10*(3), 261-266.

Forget, H., Lacroix, A., & Cohen, H. (2002). Persistent cognitive impairment following surgical treatment of Cushing's syndrome. *Psychoneuroendocrinology, 27*(3), 367-383.

Forget, H., Lacroix, A., Somma, M., & Cohen, H. (2000). Cognitive decline in patients with Cushing's syndrome. *J Int Neuropsychol Soc, 6*(1), 20-29.

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*(2), 137-144.

Friess, E., Schiffelholz, T., Steckler, T., & Steiger, A. (2000). Dehydroepiandrosterone--a neurosteroid. *Eur J Clin Invest, 30 Suppl 3*, 46-50.

Gallagher, P., Leitch, M. M., Massey, A. E., McAllister-Williams, R. H., & Young, A. H. (2006). Assessment of cortisol and dehydroepiandrosterone (DHEA) in saliva: effects of collection methods. *Journal of Psychopharmacology, 20*(5), 643-649.

Gallagher, P., Malik, N., Newham, J., Young, A. H., Ferrier, I. N., & Mackin, P. (2008). Antiglucocorticoid treatments for mood disorders. *Cochrane Database of Systematic Reviews*(1), CD005168.

Gallagher, P., Reid, K. S., & Ferrier, I. N. (2009). Neuropsychological functioning in health and mood disorder: Modulation by glucocorticoids and their receptors. *Psychoneuroendocrinology, 34* (Suppl 1), 196-207.

Gallagher, P., Watson, S., Smith, M. S., Young, A. H., & Ferrier, I. N. (2007). Plasma cortisol-dehydroepiandrosterone (DHEA) ratios in schizophrenia and bipolar disorder. *Schizophrenia Research, 90*(1-3), 258-265.

Gallagher, P., & Young, A. (2002). Cortisol/DHEA ratios in depression. *Neuropsychopharmacology, 26*(3), 410.

Garde, A. H., Hansen, A. M., Skovgaard, L. T., & Christensen, J. M. (2000). Seasonal and biological variation of blood concentrations of total cholesterol, dehydroepiandrosterone sulfate, hemoglobin A1c, IgA, prolactin, and free testosterone in healthy women. *Clinical Chemistry, 46*(4), 551-559.

Gartside, S. E., Leitch, M. M., & Young, A. H. (2003). Altered glucocorticoid rhythm attenuates the ability of a chronic SSRI to elevate forebrain 5-HT: implications for the treatment of depression. *Neuropsychopharmacology, 28*(9), 1572-1578.

Geary, D. C. (1995). Sexual selection and sex differences in spatial cognition. *Learning and Individual Differences, 7*(4), 289-301.

Gibbons, J. L. (1964). Cortisol secretion rate in depressive illness. *Arch Gen Psychiatry, 10*, 572–575.

Gibbons, J. L. (1966). The secretion rate of corticosterone in depressive illness. *Journal of Psychosomatic Research, 10*(3), 263-266.

Gibbons, J. L., & McHugh, P. R. (1962). Plasma cortisol in depressive illness. *Journal of Psychiatric Research, 2*(1), 162-171.

Girdler, S. S., & Klatzkin, R. (2007). Neurosteroids in the context of stress: implications for depressive disorders. *Pharmacol Ther, 116*(1), 125-139.

Glahn, D. C., Bearden, C. E., Cakir, S., Barrett, J. A., Najt, P., Monkul, E. S., Maples, N., Velligan, D. I., & Soares, J. C. (2006). Differential working memory impairment in bipolar disorder and schizophrenia: effects of lifetime history of psychosis. *Bipolar Disorders, 8*(2), 117-123.

Gleissner, U., & Elger, C. E. (2001). The hippocampal contribution to verbal fluency in patients with temporal lobe epilepsy. *Cortex, 37*(1), 55-63.

Gomez, R. G., Fleming, S. H., Keller, J., Flores, B., Kenna, H., DeBattista, C., Solvason, B., & Schatzberg, A. F. (2006). The neuropsychological profile of psychotic major depression and its relation to cortisol. *Biological Psychiatry, 60*(5), 472-478.

Goodwin, F. K., & Jamison, K. R. (2007). *Manic-Depressive Illness: Bipolar Disorders and Recurrent Depression* (2nd ed.). USA: OUP.

Goodyer, I. M., Herbert, J., & Altham, P. M. (1998). Adrenal steroid secretion and major depression in 8- to 16-year-olds, III. Influence of cortisol/DHEA ratio at presentation on subsequent rates of disappointing life events and persistent major depression. *Psychological Medicine, 28*(2), 265-273.

Goodyer, I. M., Herbert, J., Altham, P. M., Pearson, J., Secher, S. M., & Shiers, H. M. (1996). Adrenal secretion during major depression in 8- to 16-year-olds, I. Altered diurnal rhythms in salivary cortisol and dehydroepiandrosterone (DHEA) at presentation. *Psychol Med, 26*(2), 245-256.

Goodyer, I. M., Herbert, J., & Tamplin, A. (2003). Psychoendocrine antecedents of persistent first-episode major depression in adolescents: a community-based longitudinal enquiry.[see comment]. *Psychological Medicine, 33*(4), 601-610.

Goodyer, I. M., Herbert, J., Tamplin, A., & Altham, P. M. (2000a). First-episode major depression in adolescents. Affective, cognitive and endocrine characteristics of risk status and predictors of onset. *British Journal of Psychiatry, 176*, 142-149.

Goodyer, I. M., Herbert, J., Tamplin, A., & Altham, P. M. (2000b). Recent life events, cortisol, dehydroepiandrosterone and the onset of major depression in high-risk adolescents. *British Journal of Psychiatry, 177*, 499-504.

Goodyer, I. M., Park, R. J., & Herbert, J. (2001a). Psychosocial and endocrine features of chronic first-episode major depression in 8-16 year olds. *Biological Psychiatry, 50*(5), 351-357.

Goodyer, I. M., Park, R. J., Netherton, C. M., & Herbert, J. (2001b). Possible role of cortisol and dehydroepiandrosterone in human development and psychopathology. *British Journal of Psychiatry, 179*, 243-249.

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. *Am J Psychiatry, 165*(6), 731-739.

Goswami, U., Sharma, A., Varma, A., Gulrajani, C., Ferrier, I. N., Young, A. H., Gallagher, P., Thompson, J. M., & Moore, P. B. (2009). The neurocognitive performance of drug-free and medicated euthymic bipolar patients do not differ. *Acta Psychiatrica Scandinavica, 120*(6), 456-463.

Granger, D. A., Schwartz, E. B., Booth, A., Curran, M., & Zakaria, D. (1999). Assessing dehydroepiandrosterone in saliva: a simple radioimmunoassay for use in studies of children, adolescents and adults. *Psychoneuroendocrinology, 24*(5), 567-579.

Grunberg, S. M., Weiss, M. H., Spitz, I. M., Ahmadi, J., Sadun, A., Russel, C. A., Lucci, L., & Stevenson, L. L. (1991). Treatment of unresectable meningiomas with the antiprogesterone agent mifepristone. *J. NEUROSURG., 74*(6), 861-866.

Guazzo, E. P., Kirkpatrick, P. J., Goodyer, I. M., Shiers, H. M., & Herbert, J. (1996). Cortisol, dehydroepiandrosterone (DHEA), and DHEA sulfate in the cerebrospinal fluid of man: relation to blood levels and the effects of age. *J Clin Endocrinol Metab, 81*(11), 3951-3960.

Gurnell, E. M., & Chatterjee, V. K. K. (2001). Dehydroepiandrosterone replacement therapy. *European Journal of Endocrinology, 145*(2), 103-106.

Haitovsky, Y. (1969). Multicollinearity in Regression Analysis: Comment *The Review of Economics and Statistics, 51*(4), 486-489.

Halbreich, U., Zumoff, B., Kream, J., & Fukushima, D. K. (1982). The mean 1300-1600 h plasma cortisol concentration as a diagnostic test for hypercortisolism. *J Clin Endocrinol Metab, 54*(6), 1262-1264.

Halpern, D. F., & Wright, T. M. (1996). A process-oriented model of cognitive sex differences. *Learning and Individual Differences, 8*(1), 3-24.

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

Hansen, A. M., Garde, A. H., Skovgaard, L. T., & Christensen, J. M. (2001). Seasonal and biological variation of urinary epinephrine, norepinephrine, and cortisol in healthy women. *Clin Chim Acta, 309*(1), 25-35.

Hansen, S. L., Fjalland, B., & Jackson, M. B. (1999). Differential blockade of gamma - aminobutyric acid type A receptors by the neuroactive steroid dehydroepiandrosterone sulfate in posterior and intermediate pituitary. *Mol Pharmacol, 55*(3), 489-496.

Hatzinger, M., Hemmeter, U. M., Baumann, K., Brand, S., & Holsboer-Trachsler, E. (2002). The combined DEX-CRH test in treatment course and long-term outcome of major depression. *Journal of Psychiatric Research, 36*(5), 287-297.

Hazra, B. G., & Pore, V. S. (2001). Mifepristone (RU-486), the recently developed antiprogesterone drug and its analogues. *J. Indian Inst. Sci., 81*, 287–298.

Hechter, O., Grossman, A., & Chatterton, R. T. (1997). Relationship of dehydroepiandrosterone and cortisol in disease. *Med Hypotheses, 49*(1), 85-91.

Heikinheimo, O., Kekkonen, R., & Lahteenmaki, P. (2003a). The pharmacokinetics of mifepristone in humans reveal insights into differential mechanisms of antiprogestin action. *Contraception, 68*(6), 421-426.

Heikinheimo, O., Raivio, T., Honkanen, H., Ranta, S., & Janne, O. A. (2003b). Termination of pregnancy with mifepristone and prostaglandin suppresses transiently circulating glucocorticoid bioactivity. *J Clin Endocrinol Metab, 88*(1), 323-326.

Herbert, J., Goodyer, I. M., Altham, P. M., Pearson, J., Secher, S. M., & Shiers, H. M. (1996). Adrenal secretion and major depression in 8- to 16-year-olds, II. Influence of co-morbidity at presentation. *Psychol Med, 26*(2), 257-263.

Herrmann, W., Wyss, R., Riondel, A., Philibert, D., Teutsch, G., Sakiz, E., & Baulieu, E. E. (1982). The effects of an antiprogesterone steroid in women: interruption of the menstrual cycle and of early pregnancy. *C R Seances Acad Sci III, 294*(18), 933-938.

Het, S., Ramlow, G., & Wolf, O. T. (2005). A meta-analytic review of the effects of acute cortisol administration on human memory. *Psychoneuroendocrinology, 30*(8), 771-784.

Heuser, I., Deuschle, M., Luppa, P., Schweiger, U., Standhardt, H., & Weber, B. (1998). Increased diurnal plasma concentrations of dehydroepiandrosterone in depressed patients. *J Clin Endocrinol Metab, 83*(9), 3130-3133.

Heuser, I., Yassouridis, A., & Holsboer, F. (1994). The combined dexamethasone/CRH test: a refined laboratory test for psychiatric disorders. *J Psychiatr Res, 28*(4), 341-356.

Hinkley, L. B. N., Krubitzer, L. A., Padberg, J., & Disbrow, E. A. (2009). Visual-manual exploration and posterior parietal cortex in humans. *Journal of Neurophysiology, 102*(6), 3433-3446.

Holmes, M. K., Erickson, K., Luckenbaugh, D. A., Drevets, W. C., Bain, E. E., Cannon, D. M., Snow, J., Sahakian, B. J., Manji, H. K., & Jr, C. A. Z. (2008). A comparison of cognitive functioning in medicated and unmedicated subjects with bipolar depression. *Bipolar Disorders, 10*(7), 806-815.

Hucklebridge, F., Hussain, T., Evans, P., & Clow, A. (2005). The diurnal patterns of the adrenal steroids cortisol and dehydroepiandrosterone (DHEA) in relation to awakening. *Psychoneuroendocrinology, 30*(1), 51-57.

Ising, M., Horstmann, S., Kloiber, S., Lucae, S., Binder, E. B., Kern, N., Kunzel, H. E., Pfennig, A., Uhr, M., & Holsboer, F. (2007). Combined dexamethasone/Corticotropin Releasing Hormone test predicts treatment response in major depression-a potential biomarker? *Biological Psychiatry, 62*(1), 47-54.

Jager, G., & Postma, A. (2003). On the hemispheric specialization for categorical and coordinate spatial relations: a review of the current evidence. *Neuropsychologia, 41*(4), 504-515.

Johnson, M. D., Bebb, R. A., & Sirrs, S. M. (2002). Uses of DHEA in aging and other disease states. *Ageing Res Rev, 1*(1), 29-41.

Jozuka, H., Jozuka, E., Takeuchi, S., & Nishikaze, O. (2003). Comparison of immunological and endocrinological markers associated with major depression. *J Int Med Res, 31*(1), 36-41.

Judd, L. L., Akiskal, H. S., Maser, J. D., Zeller, P. J., Endicott, J., Coryell, W., Paulus, M. P., Kunovac, J. L., Leon, A. C., Mueller, T. I., Rice, J. A., & Keller, M. B. (1998). Major

depressive disorder: a prospective study of residual subthreshold depressive symptoms as predictor of rapid relapse. *J Affect Disord, 50*(2-3), 97-108.

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. *Arch Gen Psychiatry, 59*(6), 530-537.

Jung-Testas, I., & Baulieu, E. E. (1983). Inhibition of glucocorticosteroid action in cultured L-929 mouse fibroblasts by RU 486, a new anti-glucocorticosteroid of high affinity for the glucocorticosteroid receptor. *Exp Cell Res, 147*(1), 177-182.

Kahl, K. G., Bens, S., Ziegler, K., Rudolf, S., Dibbelt, L., Kordon, A., & Schweiger, U. (2006). Cortisol, the cortisol-dehydroepiandrosterone ratio, and pro-inflammatory cytokines in patients with current major depressive disorder comorbid with borderline personality disorder. *Biological Psychiatry, 59*(7), 667-671.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141-151.

Kalimi, M., Shafagoj, Y., Loria, R., Padgett, D., & Regelson, W. (1994). Anti-glucocorticoid effects of dehydroepiandrosterone (DHEA). *Mol Cell Biochem, 131*(2), 99-104.

Kass, R. A., & Tinsley, H. E. A. (1979). Factor analysis. *Journal of Leisure Research 11*(2), 120-138.

Kathol, R. G. (1985). Persistent elevation of urinary free cortisol and loss of circannual periodicity in recovered depressive patients. *J Affect Disord., 8*(2), 137-145.

Kathol, R. G., Anton, R., Noyes, R., & Gehris, T. (1989). Direct comparison of urinary free cortisol excretion in patients with depression and panic disorder. *Biological Psychiatry, 25*(7), 873-878.

Kelly, W. F., Checkley, S. A., Bender, D. A., & Mashiter, K. (1983). Cushing's syndrome and depression--a prospective study of 26 patients. *Br J Psychiatry, 142*, 16-19.

Kemmerer, D. (2006). The semantics of space: Integrating linguistic typology and cognitive neuroscience. *Neuropsychologia, 44*(9), 1607-1621.

Kemmerer, D., & Tranel, D. (2000). A double dissociation between linguistic and perceptual representations of spatial relationships. *Cognitive Neuropsychology, 17*(5), 393 - 414.

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.

Kessels, R. P., de Haan, E. H., Kappelle, L. J., & Postma, A. (2001). Varieties of human spatial memory: a meta-analysis on the effects of hippocampal lesions. *Brain Res Brain Res Rev, 35*(3), 295-303.

Kessels, R. P., de Haan, E. H., Kappelle, L. J., & Postma, A. (2002). Selective impairments in spatial memory after ischaemic stroke. *J Clin Exp Neuropsychol, 24*(1), 115-129.

Kessels, R. P., Hendriks, M., Schouten, J., Van Asselen, M., & Postma, A. (2004). Spatial memory deficits in patients after unilateral selective amygdalohippocampectomy. *J Int Neuropsychol Soc, 10*(6), 907-912.

Kessels, R. P. C., & Postma, A. (2002). Verbal interference during encoding and maintenance of spatial information in working memory. *Current Psychology Letters, 3*, 9.

Kessels, R. P. C., Postma, A., & de Haan, E. H. F. (1999). Object Relocation: A program for setting up, running, and analyzing experiments on memory for object locations. *Behavior Research Methods, Instruments, & Computers, 31*(3), 423-428.

Kessels, R. P. C., Postma, A., Kappelle, L. J., & de Haan, E. H. F. (2000a). Spatial memory impairment in patients after tumour resection: evidence for a double dissociation. *Journal of Neurology, Neurosurgery and Psychiatry, 69*, 389-391.

Kessels, R. P. C., Postma, A., Wester, A. J., & de Haan, E. H. F. (2000b). Memory for Object Locations in Korsakoff's Amnesia. *Cortex, 36*(1), 47-57.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H. U., & Kendler, K. S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry, 51*(1), 8-19.

King, J. A., Rosal, M. C., Ma, Y., Reed, G., Kelly, T. A., Stanek, E. J., 3rd, & Ockene, I. S. (2000). Sequence and seasonal effects of salivary cortisol. *Behav Med, 26*(2), 67-73.

Kirschbaum, C., & Hellhammer, D. H. (1989). Salivary cortisol in psychobiological research: an overview. *Neuropsychobiology, 22*(3), 150-169.

Kirschbaum, C., & Hellhammer, D. H. (1994). Salivary cortisol in psychoneuroendocrine research: recent developments and applications. *Psychoneuroendocrinology, 19*(4), 313-333.

Kirschbaum, C., & Hellhammer, D. H. (2000). Salivary cortisol. In G. Fink (Ed.), *Encyclopedia of Stress*. San Diego: Academic Press.

Kleist, K. (1911). Die klinische Stellung der Motilitätspsychosen (Vortrag auf der Versammlung des Vereins bayerischer Psychiater, München, 6.-7-6-1911). *Z. Gesamte Neurol. Psychiatr. Referate, 3*, 914–977.

Kling, M. A., Whitfield, H. J. J., Brandt, H. A., Demitrack, M. A., Kalogeras, K., Geracioti, T. D., Perini, G. I., Calabrese, J. R., Chrousos, G. P., & Gold, P. W. (1989). Effects of glucocorticoid antagonism with RU 486 on pituitary-adrenal function in patients with major depression: time-dependent enhancement of plasma ACTH secretion. *Psychopharmacology Bulletin, 25*(3), 466-472.

Knable, M. B., Torrey, E. F., Webster, M. J., & Bartko, J. J. (2001). Multivariate analysis of prefrontal cortical data from the Stanley Foundation Neuropathology Consortium. *Brain Res Bull, 55*(5), 651-659.

Knorr, U., Vinberg, M., Kessing, L. V., & Wetterslev, J. (2010). Salivary cortisol in depressed patients versus control persons: A systematic review and meta-analysis. *Psychoneuroendocrinology, 35*(9), 1275-1286.

Kosanke, J., & Bergstralh, E. (1995). SAS Match algorithm. Mayo Clinic, Division of Biostatistics: http://mayoresearch.mayo.edu/mayo/research/biostat/sasmacros.cfm.

Kosslyn, S. M., Chabris, C. F., Marsolek, C. J., & Koenig, O. (1992). Categorical versus coordinate spatial relations: computational analyses and computer simulations. *Journal of Experimental Psychology: Human Perception & Performance, 18*(2), 562-577.

Kosslyn, S. M., Koenig, O., Barrett, A., Cave, C. B., Tang, J., & Gabrieli, J. D. E. (1989). Evidence for two types of spatial representations: hemispheric specialization for categorical and coordinate relations. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 723-735.

Krishnan, K. R., Reed, D., Wilson, W. H., Saunders, W. B., Ritchie, J. C., Nemeroff, C. B., & Carroll, B. J. (1992). RU486 in depression. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 16*(6), 913-920.

Kroboth, P. D., Amico, J. A., Stone, R. A., Folan, M., Frye, R. F., Kroboth, F. J., Bigos, K. L., Fabian, T. J., Linares, A. M., Pollock, B. G., & Hakala, C. (2003). Influence of DHEA administration on 24-hour cortisol concentrations. *Journal of Clinical Psychopharmacology, 23*(1), 96-99.

Kuhlmann, S., & Wolf, O. T. (2005). Cortisol and memory retrieval in women: influence of menstrual cycle and oral contraceptives. *Psychopharmacology, 183*(1), 65-71.

Kurtz, M. M., & Gerraty, R. T. (2009). A meta-analytic investigation of neurocognitive deficits in bipolar illness: Profile and effects of clinical state. *Neuropsychology Review, 23*(5), 551-562.

Lac, G. (2001). Saliva assays in clinical and research biology. *Pathologie Biologie, 49*(8), 660-667.

Lahteenmaki, P., Heikinheimo, O., Croxatto, H., Spitz, I., Shoupe, D., Birgerson, L., & Luukkainen, T. (1987). Pharmacokinetics and metabolism of RU 486. *J Steroid Biochem, 27*(4-6), 859-863.

Lanfumey, L., Pardon, M.-C., Laaris, N., Joubert, C., Hanoun, N., Hamon, M., & Cohen-Salmon, C. (1999). 5-HT1A autoreceptor desensitization by chronic ultramild stress in mice. *Neuroreport, 10*(16), 3369-3374.

Laue, L., Gallucci, W., Loriaux, D. L., Udelsman, R., & Chrousos, G. P. (1988). The antiglucocorticoid and antiprogestin steroid RU 486: its glucocorticoid agonist effect is inadequate to prevent adrenal insufficiency in primates. *J Clin Endocrinol Metab., 67*(3), 602-606.

Laue, L., Lotze, M. T., Chrousos, G. P., Barnes, K., Loriaux, D. L., & Fleisher, T. A. (1990). Effect of chronic treatment with the glucocorticoid antagonist RU 486 in man: toxicity, immunological, and hormonal aspects. *J Clin Endocrinol Metab, 71*(6), 1474-1480.

Leitch, M. M., Ingram, C. D., Young, A. H., McQuade, R., & Gartside, S. E. (2003). Flattening the corticosterone rhythm attenuates 5-HT(1A) autoreceptor function in the rat: relevance for depression. *Neuropsychopharmacology, 28*(1), 119-125.

Leminen, R., Ranta, S., von Hertzen, H., Oehler, J., & Heikinheimo, O. (2003). Pharmacokinetics of 10 mg of mifepristone. *Contraception, 68*(6), 427-429.

Levine, A., Zagoory-Sharon, O., Feldman, R., Lewis, J. G., & Weller, A. (2007). Measuring cortisol in human psychobiological studies. *Physiology & Behavior, 90*(1), 43-53.

Levy, B., & Weiss, R. D. (2010). Neurocognitive impairment and psychosis in bipolar I disorder during early remission from an acute episode of mood disturbance. *J Clin Psychiatry, 71*(2).

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (Fourth ed.). New York: 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., & Ae, S. O. P. s. t. (2005). Incidence of bipolar affective disorder in three UK cities: Results from the AeSOP study. *The British Journal of Psychiatry, 186*(2), 126-131.

Lopez, J. F., Little, K. Y., Lopez-Figueroa, A. L., Lopez-Figueroa, M. O., & Watson, S. J. (2003). Glucocorticoid and mineralocorticoid receptor mRNA levels in the hippocampus and prefrontal cortex of subjects with mood disorders and schizophrenia. *Biological Psychiatry, 53 (suppl. 1)*(8), S173.

Lupien, S. J., Gillin, C. J., & Hauger, R. L. (1999). Working memory is more sensitive than declarative memory to the acute effects of corticosteroids: a dose-response study in humans. *Behavioral Neuroscience, 113*(3), 420-430.

Lupien, S. J., & McEwen, B. S. (1997). The acute effects of corticosteroids on cognition: integration of animal and human model studies. *Brain Research Reviews, 24*(1), 1-27.

MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences, 4*(10), 383-391.

Maes, M., Lin, A., Bonaccorso, S., van Hunsel, F., Van Gastel, A., Delmeire, L., Biondi, M., Bosmans, E., Kenis, G., & Scharpé, S. (1998). Increased 24-hour urinary cortisol excretion in patients with post-traumatic stress disorder and patients with major depression, but not in patients with fibromyalgia. *Acta Psychiatr Scand, 98*(4), 328-335.

Malhi, G. S., Ivanovski, B., Szekeres, V., & Olley, A. (2004). Bipolar disorder: It's all in your mind? The neuropsychological profile of a biological disorder. *Can J Psychiatry, 49*, 813-819.

Mannie, Z. N., Harmer, C. J., & Cowen, P. J. (2007). Increased waking salivary cortisol levels in young people at familial risk of depression. *Am J Psychiatry, 164*(4), 617-621.

Mao, X., & Barger, S. W. (1998). Neuroprotection by dehydroepiandrosterone-sulfate: role of an NFkappaB-like factor. *Neuroreport, 9*(4), 759-763.

Martinez-Aran, A., Vieta, E., Colom, F., Reinares, M., Benabarre, A., Gasto, C., & Salamero, M. (2000). Cognitive dysfunctions in bipolar disorder: Evidence of neuropsychological disturbances. *Psychotherapy and Psychosomatics, 69*(1), 2-18.

Martinez-Aran, A., Vieta, E., Colom, F., Reinares, M., Benabarre, A., Torrent, C., Goikolea, J. M., Corbella, B., Sanchez-Moreno, J., & Salamero, M. (2002). Neuropsychological performance in depressed and euthymic bipolar patients. *Neuropsychobiology, 46*(Suppl 1), 16-21.

Martinez-Aran, A., Vieta, E., Reinares, M., Colom, F., Torrent, C., Sanchez-Moreno, J., Benabarre, A., Goikolea, J. M., Comes, M., & Salamero, M. (2004). Cognitive function across manic or hypomanic, depressed, and euthymic states in bipolar disorder. *Am J Psychiatry, 161*(2), 262-270.

Marx, C. E., Stevens, R. D., Shampine, L. J., Uzunova, V., Trost, W. T., Butterfield, M. I., Massing, M. W., Hamer, R. M., Morrow, A. L., & Lieberman, J. A. (2006). Neuroactive steroids are altered in schizophrenia and bipolar disorder: relevance to pathophysiology and therapeutics. *Neuropsychopharmacology, 31*, 1249-1263.

Mauri, M., Sinforiani, E., Bono, G., Vignati, F., Berselli, M. E., Attanasio, R., & Nappi, G. (1993). Memory impairment in Cushing's disease. *Acta Neurol Scand, 87*(1), 52-55.

Maurice, T., Gregoire, C., & Espallergues, J. (2006). Neuro(active)steroids actions at the neuromodulatory sigma1 ([sigma]1) receptor: Biochemical and physiological

evidences, consequences in neuroprotection. *Pharmacology Biochemistry and Behavior, 84*(4), 581-597.

Maurice, T., Phan, V. L., Urani, A., Kamei, H., Noda, Y., & Nabeshima, T. (1999). Neuroactive neurosteroids as endogenous effectors for the sigma1 ([sigma]1) receptor: pharmacological evidence and therapeutic opportunities. *Jpn. J. Pharmacol, 81*(2), 125-155.

McAllister-Williams, R. H., Ferrier, I. N., & Young, A. H. (1998). Mood and neuropsychological function in depression: the role of corticosteroids and serotonin. *Psychological Medicine, 28*(3), 573-584.

McEwen, B. S. (2010). Stress, sex, and neural adaptation to a changing environment: mechanisms of neuronal remodeling. *Ann N Y Acad Sci, 1204*(Suppl), E38-59.

McNamara, T. P. (1991). Memory's view of space. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 27). London: Academic Press.

McNamara, T. P., Halpin, J. A., & Hardy, J. K. (1992). Spatial and temporal contributions to the structure of spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(3), 555-564.

Michael, A., Jenaway A, Paykel, E. S., & Herbert, J. (2000). Altered salivary dehydroepiandrosterone levels in major depression in adults. *Biological Psychiatry, 48*(10), 989-995.

Modell, S., Yassouridis, A., Huber, J., & Holsboer, F. (1997). Corticosteroid receptor function is decreased in depressed patients. *Neuroendocrinology, 65*(3), 216-222.

Mojtabai, R., Bromet, E. J., Harvey, P. D., Carlson, G. A., Craig, T. J., & Fennig, S. (2000). Neuropsychological differences between first-admission schizophrenia and psychotic affective disorders. *American Journal of Psychiatry, 157*, 1453-1460.

Montgomery, S. A., & Asberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry, 134*, 382-389.

Morfin, R. (2002). *DHEA and the brain*. London and New York: Taylor and Francis.

Morris, R. G., Pickering, A., Abrahams, S., & Feigenbaum, J. D. (1996). Space and the hippocampal formation in humans. *Brain Research Bulletin, 40*(5-6), 487-490.

Muller, C., Hennebert, O., & Morfin, R. (2006). The native anti-glucocorticoid paradigm. *The Journal of Steroid Biochemistry and Molecular Biology, 100*(1-3), 95-105.

Murphy, B. E., Filipini, D., & Ghadirian, A. M. (1993). Possible use of glucocorticoid receptor antagonists in the treatment of major depression: preliminary results using RU 486. *Journal of Psychiatry & Neuroscience, 18*(5), 209-213.

Nelson, H. E. (1982). *National Adult Reading Test, NART.* Windsor: Nelson Publishing Company.

Nelson, J. C., & Davis, J. M. (1997). DST studies in psychotic depression: a meta-analysis. *American Journal of Psychiatry, 154*(11), 1497-1503.

Neu, P., Kiesslinger, U., Schlattmann, P., & Reischies, F. M. (2001). Time-related cognitive deficiency in four different types of depression. *Psychiatry Research, 103*(2-3), 237-247.

Newcomer, J. W., Selke, G., Melson, A. K., Hershey, T., Craft, S., K., R., & Alderson, A. L. (1999). Decreased memory performance in healthy humans induced by stress-level cortisol treatment. *Arch Gen Psychiatry, 56*(6), 527-533.

Ng, B., Camacho, A., Lara, D. R., Brunstein, M. G., Pinto, O. C., & Akiskal, H. S. (2008). A case series on the hypothesized connection between dementia and bipolar spectrum disorders: Bipolar type VI? *Journal of Affective Disorders, 107*(1-3), 307-315.

Nieman, L. K., Chrousos, G. P., Kellner, C., Spitz, I. M., Nisula, B. C., Cutler, G. B., Merriam, G. R., Bardin, C. W., & Loriaux, D. L. (1985). Successful treatment of Cushing's syndrome with the glucocorticoid antagonist RU 486. *Journal of Clinical Endocrinology & Metabolism, 61*(3), 536-540.

Oberbeck, R., Benschop, R. J., Jacobs, R., Hosch, W., Jetschmann, J. U., Schurmeyer, T. H., Schmidt, R. E., & Schedlowski, M. (1998). Endocrine mechanisms of stress-induced DHEA-secretion. *J Endocrinol Invest, 21*(3), 148-153.

Oitzl, M. S., & de Kloet, E. R. (1992). Selective Corticosteroid Antagonists Modulate Specific Aspects of Spatial Orientation Learning. *Behavioral Neuroscience, 106*(1), 62-71.

Oitzl, M. S., Fluttert, M., & de Kloet, E. R. (1998a). Acute blockade of hippocampal glucocorticoid receptors facilitates spatial learning in rats. *Brain Research, 797*(1), 159-162.

Oitzl, M. S., Fluttert, M., Sutanto, W., & de Kloet, E. R. (1998b). Continuous blockade of brain glucocorticoid receptors facilitates spatial learning and memory in rats. *European Journal of Neuroscience, 10*(12), 3759 -3766.

Oldehinkel, A. J., van den Berg, M. D., Flentge, F., Bouhuys, A. L., ter Horst, G. J., & Ormel, J. (2001). Urinary free cortisol excretion in elderly persons with minor and major depression. *Psychiatry Research, 104*(1), 39-47.

Osran, H., Reist, C., Chen, C. C., Lifrak, E. T., Chicz-DeMet, A., & Parker, L. N. (1993). Adrenal androgens and cortisol in major depression. *American Journal of Psychiatry, 150*(5), 806-809.

Overall, J. E., & Gorham, D. R. (1962). The Brief Psychiatric Rating Scale. *Psychological Reports, 10*, 799-812.

Owen, A. M., Downes, J. J., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1990). Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia, 28*(10), 1021-1034.

Owen, A. M., Sahakian, B. J., Semple, J., Polkey, C. E., & Robbins, T. W. (1995). Visuo-spatial short-term recognition memory and learning after temporal lobe excisions, frontal lobe excisions or amygdalo- hippocampectomy in man. *Neuropsychologia, 33*(1), 1-24.

Pachet, A. K., & Wisniewski, A. M. (2003). The effects of lithium on cognition: an updated review. *Psychopharmacology, 170*(3), 225 - 234.

Pariante, C. M. (2003). Depression, stress and the adrenal axis. *J Neuroendocrinol, 15*(8), 811-812.

Pariante, C. M. (2004). Glucocorticoid receptor function in vitro in patients with major depression. *Stress, 7*(4), 209-219.

Pariante, C. M., Papadopoulos, A. S., Poon, L., Cleare, A. J., Checkley, S. A., English, J., Kerwin, R. W., & Lightman, S. (2004a). Four days of citalopram increase suppression of cortisol secretion by prednisolone in healthy volunteers. *Psychopharmacology (Berl), 177*(1-2), 200-206.

Pariante, C. M., Thomas, S. A., Lovestone, S., Makoff, A., & Kerwin, R. W. (2004b). Do antidepressants regulate how cortisol affects the brain? *Psychoneuroendocrinology, 29*(4), 423-447.

Parker, L. N., Levin, E. R., & Lifrak, E. T. (1985). Evidence for adrenocortical adaptation to severe illness. *J Clin Endocrinol Metab, 60*(5), 947-952.

Parrot, M., Doyon, B., Demonet, J.-F., & Cardebat, D. (1999). Hemispheric preponderance in categorical and coordinate visual processes. *Neuropsychologia, 37*(11), 1215-1225.

Patel, P. D., Lopez, J. F., Lyons, D. M., Burke, S., Wallace, M., & Schatzberg, A. F. (2000). Glucocorticoid and mineralocorticoid receptor mRNA expression in squirrel monkey brain. *J Psychiatr Res, 34*(6), 383-392.

Patil, C. G., Lad, S. P., Katznelson, L., & Laws, E. R. (2007). Brain atrophy and cognitive deficits in Cushing's disease. *Neurosurg Focus, 23*(3), E11.

Pearson, A., de Vries, A., Middleton, S., Gillies, F., White, T., Armstrong, I., Andrew, R., Seckl, J., & MacLullich, A. (2010). Cerebrospinal fluid cortisol levels are higher in patients with delirium versus controls. *BMC Research Notes, 3*(1), 33.

Pedhauzur, E., & Schmelkin, L. (1991). *Measurement, design and analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Petrides, M., & Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia, 20*(3), 249-262.

Piekema, C., Kessels, R. P., Mars, R. B., Petersson, K. M., & Fernandez, G. (2006). The right hippocampus participates in short-term memory maintenance of object-location associations. *Neuroimage, 33*(1), 374-382.

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*(4), 425-434.

Pomara, N., Doraiswamy, P. M., Tun, H., & Ferris, S. (2002). Mifepristone (RU 486) for Alzheimer's disease. *Neurology, 58*(9), 1436-1437.

Poor, V., Juricskay, S., Gati, A., Osvath, P., & Tenyi, T. (2004). Urinary steroid metabolites and 11beta-hydroxysteroid dehydrogenase activity in patients with unipolar recurrent major depression. *Journal of Affective Disorders, 81*(1), 55-59.

Popescu, C., Ionescu, R., Jipescu, I., & Popa, S. (1991). Psychomotor functioning in unipolar and bipolar affective disorders. *Rom J Neurol Psychiatry, 29*(1-2), 17-33.

Porter, R. J., Bourke, C., & Gallagher, P. (2007). Neuropsychological impairment in major depression - its nature, origin and clinical significance. *Australian and New Zealand Journal of Psychiatry, 41*, 115-128.

Porter, R. J., Gallagher, P., Watson, S., Lunn, B. S., & Young, A. H. (2002). Effects of sub-chronic administration of hydrocortisone on hormonal and psychological responses to l-tryptophan in normal male volunteers. *Psychopharmacology, 163*(1), 68-75.

Porter, R. J., Gallagher, P., Watson, S., & Young, A. H. (2004). Corticosteroid-serotonin interactions in depression: a review of the human evidence. *Psychopharmacology (Berl), 173*(1-2), 1-17.

Postma, A., & de Haan, E. H. (1996). What was where? Memory for object locations. *Quarterly Journal of Experimental Psychology A, 49*(1), 178-199.

Postma, A., Huntjens, R. J. C., Meuwissen, M., & Laeng, B. (2006). The time course of spatial memory processing in the two hemispheres. *Neuropsychologia, 44*(10), 1914-1918.

Postma, A., Izendoorn, R., & de Haan, E. H. F. (1998). Sex differences in object location memory. *Brain and Cognition, 36*(3), 334-345.

Postma, A., Kessels, R. P. C., & van Asselen, M. (2008). How the brain remembers and forgets where things are: The neurocognition of object-location memory. *Neuroscience and Biobehavioral Reviews, 32*(8), 1339-1345.

Postma, A., Meyer, G., Tuiten, A., van Honk, J., Kessels, R. P. C., & Thijssen, J. (2000). Effects of testosterone administration on selective aspects of object-location memory in healthy young women. *Psychoneuroendocrinology, 25*(6), 563-575.

Postma, A., Winkel, J., Tuiten, A., & van Honk, J. (1999). Sex differences and menstrual cycle effects in human spatial memory. *Psychoneuroendocrinology, 24*(2), 175-192.

Quirarte, G. L., Roozendaal, B., & McGaugh, J. L. (1997). Glucocorticoid enhancement of memory storage involves noradrenergic activation in the basolateral amygdala. *Proceedings of the National Academy of Sciences of the United States of America, 94*(25), 14048-14053.

Quraishi, S., & Frangou, S. (2002). Neuropsychology of bipolar disorder: a review. *J Affect Disord, 72*(3), 209-226.

Raison, C. L., & Miller, A. H. (2003). When not enough is too much: the role of insufficient glucocorticoid signaling in the pathophysiology of stress-related disorders. *Am J Psychiatry, 160*(9), 1554-1565.

Rapport, L. J., Webster, J. S., & Dutra, R. L. (1994). Digit span performance and unilateral neglect. *Neuropsychologia, 32*(5), 517-525.

Raven, P. W., & Taylor, N. F. (1998). 11Beta-HSD and 17beta-HSD as biological markers of depression: sex differences and correlation with symptom severity. *Endocr Res, 24*(3-4), 659-662.

Regelson, W., & Kalimi, M. (1994). Dehydroepiandrosterone (DHEA)–the Multifunctional Steroid II. Effects on the CNS, Cell Proliferation, Metabolic and Vascular, Clinical and Other Effects. Mechanism of Action? *Annals of the New York Academy of Sciences, 719*, 564-575.

Reiten, R. M. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual & Motor Skills, 8*, 271-276.

Reppermund, S., Zihl, J., Lucae, S., Horstmann, S., Kloiber, S., Holsboer, F., & Ising, M. (2007). Persistent cognitive impairment in depression: the role of psychopathology and altered hypothalamic-pituitary-adrenocortical (HPA) system regulation. *Biological Psychiatry, 62*(5), 400-406.

Reul, J. M., & de Kloet, E. R. (1985). Two receptor systems for corticosterone in rat brain: microdistribution and differential occupation. *Endocrinology, 117*(6), 2505-2511.

Rey, A. (1964). *L'Examen Clinique en Psychologie*. Paris: Press Universitaire de France.

Ribeiro, S. C., Tandon, R., Grunhaus, L., & Greden, J. F. (1993). The DST as a predictor of outcome in depression: a meta-analysis. *Am J Psychiatry, 150*(11), 1618-1629.

Ritsner, M. S., Gibel, A., Ratner, Y., & Weizman, A. (2008). Dehydroepiandrosterone and Pregnenolone Alterations in Schizophrenia. In M. S. Ritsner & A. Weizman (Eds.), *Neuroactive Steroids in Brain Function, Behavior and Neuropsychiatric Disorders: Novel Strategies for Research and Treatment* (pp. 251-298): Springer.

Robertson, I. H. (1990). Digit span and visual neglect: A puzzling relationship. *Neuropsychologia, 28*(2), 217-222.

Robinson, L. J., & Ferrier, I. N. (2006). 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 bipolar subjects. *Journal of Affective Disorders, 93*(1-3), 105-115.

Roiser, J. P., Cannon, D. M., Gandhi, S. K., Tavares, J. T., Erickson, K., Wood, S., Klaver, J. M., Clark, L., Zarate Jr, C. A., Sahakian, B. J., & Drevets, W. C. (2009). Hot and cold cognition in unmedicated depressed subjects with bipolar disorder. *Bipolar Disorders, 11*(2), 178-189.

Roozendaal, B., Barsegyan, A., & Lee, S. (2008). Adrenal stress hormones, amygdala activation, and memory for emotionally arousing experiences. *Prog Brain Res., 167*, 79-97.

Roozendaal, B., Okuda, S., Van der Zee, E. A., & McGaugh, J. L. (2006). Glucocorticoid enhancement of memory requires arousal-induced noradrenergic activation in the basolateral amygdala. *PNAS, 103*(17), 6741-6746.

Rosen, W. G., Mohs, R. C., & Davis, K. L. (1984). A new rating scale for Alzheimer's disease. *Am J Psychiatry, 141*(11), 1356-1364.

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.

Rush, A. J., Giles, D. E., Schlesser, M. A., Orsulak, P. J., Parker, C. R., Jr., Weissenburger, J. E., Crowley, G. T., Khatami, M., & Vasavada, N. (1996). The dexamethasone suppression test in patients with mood disorders. *J Clin Psychiatry, 57*(10), 470-484.

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*(5), 363-370.

Sartor, O., & Cutler, G. B., Jr. (1996). Mifepristone: treatment of Cushing's syndrome. *Clin Obstet Gynecol, 39*(2), 506-510.

Sauro, M. D., Jorgensen, R. S., & Teal Pedlow, C. (2003). Stress, glucocorticoids, and memory: a meta-analytic review. *Stress, 6*(4), 235-245.

Savitz, J., Solms, M., & Ramesar, R. (2005). Neuropsychological dysfunction in bipolar affective disorder: a critical opinion. *Bipolar Disorders, 7*(3), 216-235.

Schacter, D. L., & Nadel, L. (1992). Varieties of spatial memory: a problem for cognitive neuroscience. In R. G. Lister & H. J. Weingartner (Eds.), *Perspectives on cognitive neuroscience*. New York: Oxford University Press.

Schneider, J. J., Candiago, R. H., Rosa, A. R., Ceresér, K. M., & Kapczinski, F. (2008). Cognitive impairment in a Brazilian sample of patients with bipolar disorder. *Rev Bras Psiquiatr., 30*(3), 209-214.

Scott, L. V., & Dinan, T. G. (1998). Urinary free cortisol excretion in chronic fatigue syndrome, major depression and in healthy volunteers. *Journal of Affective Disorders, 47*(1-3), 49-54.

Scott, L. V., Salahuddin, F., Cooney, J., Svec, F., & Dinan, T. G. (1999). Differences in adrenal steroid profile in chronic fatigue syndrome, in depression and in health. *Journal of Affective Disorders, 54*(1-2), 129-137.

Shirtcliff, E. A., Granger, D. A., Schwartz, E., & Curran, M. J. (2001). Use of salivary biomarkers in biobehavioral research: cotton-based sample collection methods can interfere with salivary immunoassay results. *Psychoneuroendocrinology, 26*(2), 165-173.

Shores, E. A., Carstairs, J. R., & Crawford, J. R. (2006). Excluded Letter Fluency Test (ELF): Norms and test-retest reliability data for healthy young adults. *Brain Impairment, 7*(1), 26-32.

Sitruk-Ware, R., & Spitz, I. M. (2003). Pharmacological properties of mifepristone: toxicology and safety in animal and human studies. *Contraception, 68*(6), 409-420.

Smeets, T., Otgaar, H., Candel, I., & Wolf, O. T. (2008). True or false? Memory is differentially affected by stress-induced cortisol elevations and sympathetic activity at consolidation and retrieval. *Psychoneuroendocrinology, 33*(10)), 1378-1386.

Spitz, I. M., & Bardin, C. W. (1993a). Clinical pharmacology of RU 486--an antiprogestin and antiglucocorticoid. *Contraception, 48*(5), 403-444.

Spitz, I. M., & Bardin, C. W. (1993b). Mifepristone (RU 486)--a modulator of progestin and glucocorticoid action. *N Engl J Med, 329*(6), 404-412.

SPSS. (1998). *SPSS for Windows version 9*. Chicago: SPSS Inc.

SPSS. (2008). *SPSS for Windows version 17*: SPSS Inc.

Starkman, M. N., Gebarski, S. S., Berent, S., & Schteingart, D. E. (1992). Hippocampal formation volume, memory dysfunction, and cortisol levels in patients with Cushing's syndrome. *Biol Psychiatry, 32*(9), 756-765.

Starkman, M. N., Giordani, B., Berent, S., Schork, M. A., & Schteingart, D. E. (2001). Elevated cortisol levels in Cushing's disease are associated with cognitive decrements. *Psychosomatic Medicine, 63*(6), 985-993.

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4 ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary* (Third Edition ed.): OUP USA.

Sundstrom Poromaa, I., Smith, S., & Gulinello, M. (2003). GABA receptors, progesterone and premenstrual dysphoric disorder. *Arch Womens Ment Health, 6*(1), 23-41.

Suppes, T., Leverich, G. S., Keck, P. E., Nolen, W. A., Denicoff, K. D., Altshuler, L. L., McElroy, S. L., Rush, A. J., Kupka, R., Frye, M. A., Bickel, M., & Post, R. M. (2001). The Stanley Foundation Bipolar Treatment Outcome Network: II. Demographics and illness characteristics of the first 261 patients. *Journal of Affective Disorders, 67*(1-3), 45-59.

Sweeney, J. A., Kmieca, J. A., & Kupfer, D. J. (2000). Neuropsychologic impairments in bipolar and unipolar mood disorders on the CANTAB neurocognitive battery. *Biological Psychiatry, 48*(7), 674-684.

Symonds, C. S., Gallagher, P., Thompson, J. M., & Young, A. H. (2004). Effects of the menstrual cycle on mood, neurocognitive and neuroendocrine function in healthy premenopausal women. *Psychological Medicine, 34*(1), 93-102.

Takebayashi, M., Kagaya, A., Uchitomi, Y., Kugaya, A., Muraoka, M., Yokota, N., Horiguchi, J., & Yamawaki, S. (1998). Plasma dehydroepiandrosterone sulfate in unipolar major depression. *Journal of Neural Transmission, 105*(4-5), 537-542.

Taylor Tavares, J. V., Clark, L., Cannon, D. M., Erickson, K., Drevets, W. C., & Sahakian, B. J. (2007). Distinct profiles of neurocognitive function in unmedicated unipolar depression and bipolar II depression. *Biological Psychiatry, 62*(8), 917-924.

Thompson, J., Hamilton, C. J., Gray, J., Quinn, J. G., Mackin, P., Young, A., & Nicol Ferrier, I. (2006). Executive and visuospatial sketchpad resources in euthymic bipolar disorder: Implications for visuospatial working memory architecture. *Memory, 14*(4), 437-451.

Thompson, J. M., Gray, J. M., Crawford, J. R., Hughes, J. H., Young, A. H., & Ferrier, I. N. (2009). Differential deficit in executive control in euthymic bipolar disorder? *Journal of Abnormal Psychology, 118*(1), 146–160.

Tollefson, G. D., Haus, E., Garvey, M. J., Evans, M., & Tuason, V. B. (1990). 24 hour urinary dehydroepiandrosterone sulfate in unipolar depression treated with cognitive and/or pharmacotherapy. *Ann Clin Psychiatry, 2*, 39-45.

Torrent, C., Martinez-Aran, A., Daban, C., Sanchez-Moreno, J., Comes, M., Goikolea, J. M., Salamero, M., & Vieta, E. (2006). Cognitive impairment in bipolar II disorder. *Br J Psychiatry, 189*(3), 254-259.

Trenerry, M. R., Crosson, B., Deboe, J., & Leber, W. R. (1989). *Stroop Neuropsychological Screening Test*. Odessa, FL: Psychological Assessment Resources.

Tytherleigh, M. Y., Vedhara, K., & Lightman, S. L. (2004). Mineralocorticoid and glucocorticoid receptors and their differential effects on memory performance in people with Addison's disease. *Psychoneuroendocrinology, 29*(6.), 712-723.

van 't Wout, M., Aleman, A., Kessels, R. P. C., Cahn, W., de Haan, E. H. F., & Kahn, R. S. (2007). Exploring the nature of facial affect processing deficits in schizophrenia. *Psychiatry Research, 150*(3), 227-235.

van Asselen, M., Kessels, R. P. C., Frijns, C. J. M., Jaap Kappelle, L., Neggers, S. F. W., & Postma, A. (2009). Object-location memory: A lesion-behavior mapping study in stroke patients. *Brain and Cognition, 71*(3), 287-294.

van Asselen, M., Kessels, R. P. C., Kappelle, L. J., & Postma, A. (2008). Categorical and coordinate spatial representations within object-location memory. *Cortex, 44*(3), 249-256.

van Asselen, M., Kessels, R. P. C., Neggers, S. F. W., Kappelle, L. J., Frijns, C. J. M., & Postma, A. (2006). Brain areas involved in spatial working memory. *Neuropsychologia, 44*(7), 1185-1194.

van Asselen, M., Kessels, R. P. C., Wester, A. J., & Postma, A. (2005). Spatial working memory and contextual cueing in patients with Korsakoff amnesia. *Journal of Clinical and Experimental Neuropsychology, 27*, 645-655.

van Broekhoven, F., & Verkes, R. J. (2003). Neurosteroids in depression: a review. *Psychopharmacology, 165*, 97-110.

van der Ham, I. J. M., Raemaekers, M., van Wezel, R. J. A., Oleksiak, A., & Postma, A. (2009). Categorical and coordinate spatial relations in working memory: An fMRI study. *Brain Research, 1297*, 70-79.

van der Ham, I. J. M., van Wezel, R. J. A., Oleksiak, A., & Postma, A. (2007). The time course of hemispheric differences in categorical and coordinate spatial processing. *Neuropsychologia, 45*(11), 2492-2498.

van der Lely, A. J., Foeken, K., van der Mast, R. C., & Lamberts, S. W. (1991). Rapid reversal of acute psychosis in the Cushing syndrome with the cortisol-receptor antagonist mifepristone (RU 486). *Ann Intern Med, 114*(2), 143-144.

van der Lubbe, R. H. J., Scholvinck, M. L., Kenemans, J. L., & Postma, A. (2006). Divergence of categorical and coordinate spatial processing assessed with ERPs. *Neuropsychologia, 44*, 1547-1559.

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. *Arch Gen Psychiatry, 55*(1), 41-46.

van Haarst, A. D., Oitzl, M. S., & de Kloet, E. R. (1997). Facilitation of Feedback Inhibition Through Blockade of Glucocorticoid Receptors in the Hippocampus. *Neurochemical Research, 22*(11), 1323-1328.

van Londen, L., Goekoop, J. G., Zwinderman, A. H., Lanser, J. B., Wiegant, V. M., & De Wied, D. (1998). Neuropsychological performance and plasma cortisol, arginine vasopressin and oxytocin in patients with major depression. *Psychological Medicine, 28*(2), 275-284.

Vanderplas, J. M., & Garvin, E. A. (1959). The association value of random shapes. *J Exp Psychol, 57*(3), 147-154.

Vickers, A. J., & Altman, D. G. (2001). Analysing controlled trials with baseline and follow up measurements. *BMJ, 323*(10 November), 1123-1124.

Vining, R. F., McGinley, R. A., & Symons, R. G. (1983). Hormones in saliva: mode of entry and consequent implications for clinical interpretation. *Clin Chem, 29*(10), 1752-1756.

Walker, B. R., Best, R., Noon, J. P., Watt, G. C., & Webb, D. J. (1997). Seasonal variation in glucocorticoid activity in healthy men. *J Clin Endocrinol Metab, 82*(12), 4015-4019.

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*(6), 496-502.

Watson, S., Gallagher, P., Smith, M. S., Ferrier, I. N., & Young, A. H. (2006a). The dex/CRH test - is it better than the DST? *Psychoneuroendocrinology, 31*, 889-894.

Watson, S., Owen, B. M., Gallagher, P., Hearn, A. J., Young, A. H., & Ferrier, I. N. (2007). Family history, early adversity and the hypothalamic-pituitary-adrenal (HPA) axis: Mediation of the vulnerability to mood disorders. *Neuropsychiatric Disease and Treatment, 3*(5), 647-653.

Watson, S., Thompson, J. M., Ritchie, J. C., Ferrier, I. N., & Young, A. H. (2006b). Neurocognitive impairment in bipolar disorder: the relationship with glucocorticoid receptor function. *Bipolar Disorders, 8*(1), 85-90.

Wauthy, J., Ansseau, M., von Frenckell, R., Mormont, C., & Legros, J. J. (1991). Memory disturbances and dexamethasone suppression test in major depression. *Biological Psychiatry, 30*(7), 736-738.

Weber, B., Lewicka, S., Deuschle, M., Colla, M., & Heuser, I. (2000). Testosterone, androstenedione and dihydrotestosterone concentrations are elevated in female patients with major depression. *Psychoneuroendocrinology, 25*(8), 765-771.

Webster, M. J., Knable, M. B., O'Grady, J., Orthmann, J., & Weickert, C. S. (2002). Regional specificity of brain glucocorticoid receptor mRNA alterations in subjects with schizophrenia and mood disorders. *Mol Psychiatry, 7*(9), 985-994.

Wechsler, D. (1981). *WAIS-R manual, Wechsler Adult Intelligence Scale-Revised*. Cleveland, OH.: Psychological Corp.

Wedekind, D., Preiss, B., Cohrs, S., Ruether, E., Huether, G., & Adler, L. (2007). Relationship between nocturnal urinary cortisol excretion and symptom severity in subgroups of patients with depressive episodes. *Neuropsychobiology, 56*(2-3), 119-122.

Weiss, E. M., Siedentopf, C., Hofer, A., Deisenhammer, E. A., Hoptman, M. J., Kremser, C., Golaszewski, S., Felber, S., Fleischhacker, W. W., & Delazer, M. (2003). Brain activation pattern during a verbal fluency test in healthy male and female volunteers: a functional magnetic resonance imaging study. *Neuroscience Letters, 352*(3), 191-194.

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., Lépine, 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. *JAMA, 276*(4), 293-299.

Weitzman, E. D., Fukushima, D., Nogeire, C., Roffwarg, H., Gallagher, T. F., & Hellman, L. (1971). Twenty-four hour pattern of the episodic secretion of cortisol in normal subjects. *Journal of Clinical Endocrinology & Metabolism, 33*(1), 14-22.

Wen, S., Dong, K., Onolfo, J. P., & Vincens, M. (2001). Treatment with dehydroepiandrosterone sulfate increases NMDA receptors in hippocampus and cortex. *Eur J Pharmacol, 430*(2-3), 373-374.

Werner, S., & Diedrichsen, J. (2002). The time course of spatial memory distortions. *Memory and Cognition, 30*(5), 718-730.

Wernicke, C. (1900). *Grundriss der Psychiatrie* (Vol. .). Leipzig: Thieme.

Westgard, J. O., Barry, P. L., & Hunt, M. R. (1981). A multi-rule Shewhart chart for quality control in clinical chemistry. *Clinical Chemistry, 27*, 493-501.

Whelan, T. B., Schteingart, D. E., Starkman, M. N., & Smith, A. (1980). Neuropsychological deficits in Cushing's syndrome. *J Nerv Ment Dis, 168*(12), 753-757.

WHO. (1992). *ICD-10 classification of mental and behavioral disorders: Clinical descriptions and diagnostic guidelines*. Geneva: World Health Organization.

Williams, E. J. (1959). The Comparison of Regression Variables. *Journal of the Royal Statistical Society. Series B (Methodological), 21*(2), 396-399.

Wolf, O. T. (2008). The influence of stress hormones on emotional memory: relevance for psychopathology. *Acta Psychologica, 127*(3), 513-531.

Wolf, O. T., Köster, B., Kirschbaum, C., Pietrowsky, R., Kern, W., Hellhammer, D. H., Born, J., & Fehm, H. L. (1997). A single administration of dehydroepiandrosterone does not enhance memory performance in young healthy adults, but immediately reduces cortisol levels. *Biological Psychiatry, 42*(9), 845-848.

Wolf, O. T., Kuhlmann, S., Buss, C., Hellhammer, D. H., & Kirschbaum, C. (2004). Cortisol and memory retrieval in humans: influence of emotional valence. *Ann NY Acad Sci, 1032*(1), 195-197.

Wolfe, J., Granholm, E., Butters, N., Saunders, E., & Janowsky, D. (1987). Verbal memory deficits associated with major affective disorders: a comparison of unipolar and bipolar patients. *J Affect Disord, 13*(1), 83-92.

Wolkowitz, O. M., Epel, E. S., & Reus, V. I. (2001). Stress hormone-related psychopathology: pathophysiological and treatment implications. *World J Biol Psychiatry, 2*(3), 115-143.

Wolkowitz, O. M., Reus, V. I., Weingartner, H., Thompson, K., Breier, A., Doran, A., Rubinow, D., & Pickar, D. (1990). Cognitive effects of corticosteroids. *Am J Psychiatry, 147*(10), 1297-1303.

Wong, M. L., Kling, M. A., Munson, P. J., Listwak, S., Licinio, J., Prolo, P., Karp, B., McCutcheon, I. E., Geracioti, T. D., DeBellis, M. D., Rice, K. C., Goldstein, D. S., Veldhuis, J. D., Chrousos, G. P., Oldfield, E. H., McCann, S. M., & Gold, P. W. (2000). Pronounced and sustained central hypernoradrenergic function in major depression with melancholic features: relation to hypercortisolism and corticotropin-releasing hormone. *Proceedings of the National Academy of Sciences of the United States of America, 97*(1), 325-330.

Yehuda, R., Boisoneau, D., Mason, J. W., & Giller, E. L. (1993). Glucocorticoid receptor number and cortisol excretion in mood, anxiety, and psychotic disorders. *Biological Psychiatry, 34*(1-2), 18-25.

Young, A. H., Gallagher, P., & Porter, R. J. (2002). Elevation of the cortisol-dehydroepiandrosterone ratio in drug-free depressed patients. *American Journal of Psychiatry, 159*(7), 1237–1239.

Young, A. H., Sahakian, B. J., Robbins, T. W., & Cowen, P. J. (1999). The effects of chronic administration of hydrocortisone on cognitive function in normal male volunteers. *Psychopharmacology (Berl), 145*(3), 260-266.

Young, E. A., Abelson, J., & Lightman, S. L. (2004a). Cortisol pulsatility and its role in stress regulation and health. *Frontiers in Neuroendocrinology, 25*(2), 69-76.

Young, E. A., Altemus, M., Lopez, J. F., Kocsis, J. H., Schatzberg, A. F., deBattista, C., & Zubieta, J.-K. (2004b). HPA axis activation in major depression and response to fluoxetine: a pilot study. *Psychoneuroendocrinology, 29*(9), 1198-1204.

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

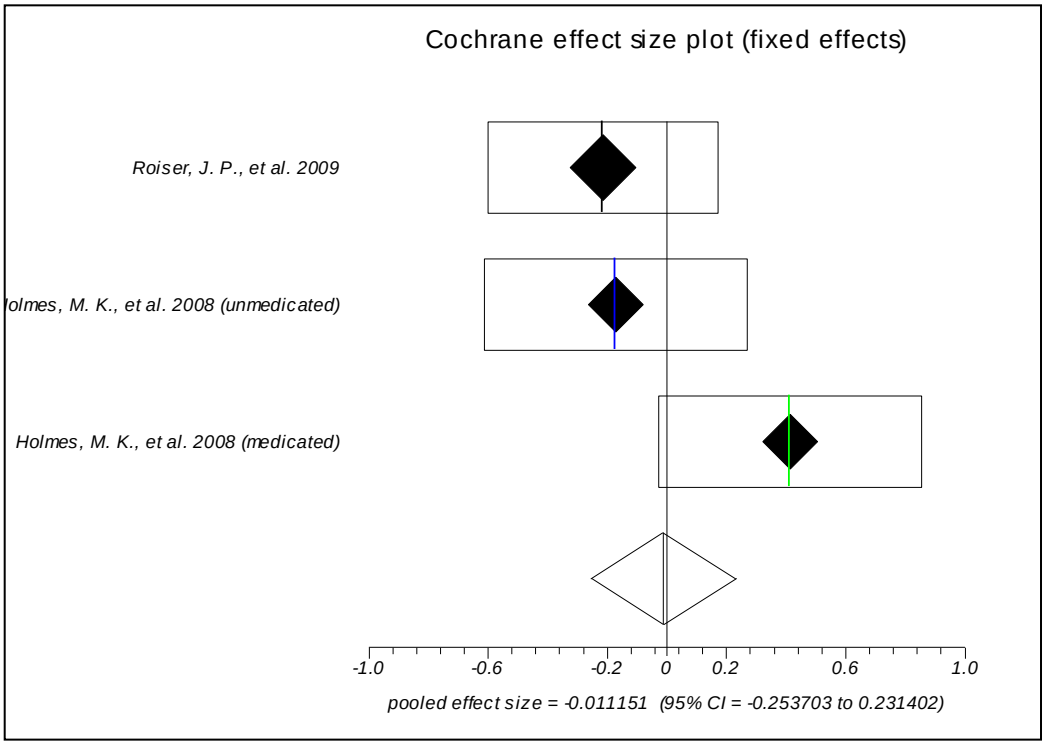
Zobel, A. W., Schulze-Rauschenbach, S., von Widdern, O. C., Metten, M., Freymann, N., Grasmader, K., Pfeiffer, U., Schnell, S., Wagner, M., & Maier, W. (2004). Improvement of working but not declarative memory is correlated with HPA normalization during antidepressant treatment. *Journal of Psychiatric Research, 38*(4), 377-383.

Appendices

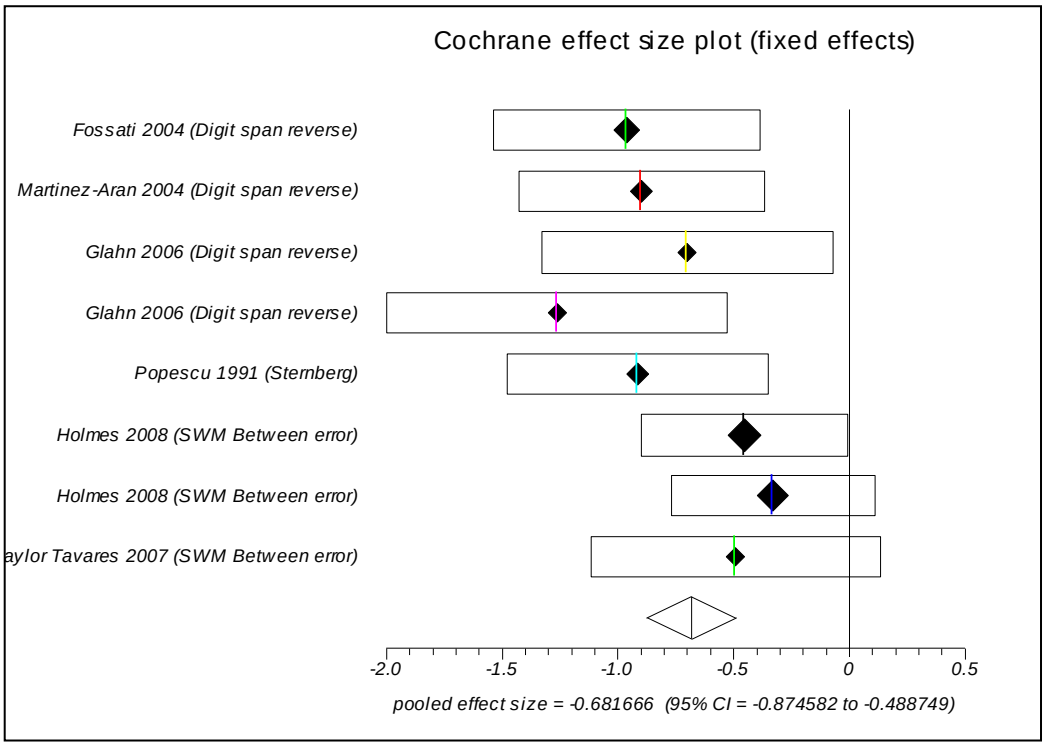
9. Appendices

9.1 Effect size plots from Chapter 1.3.2

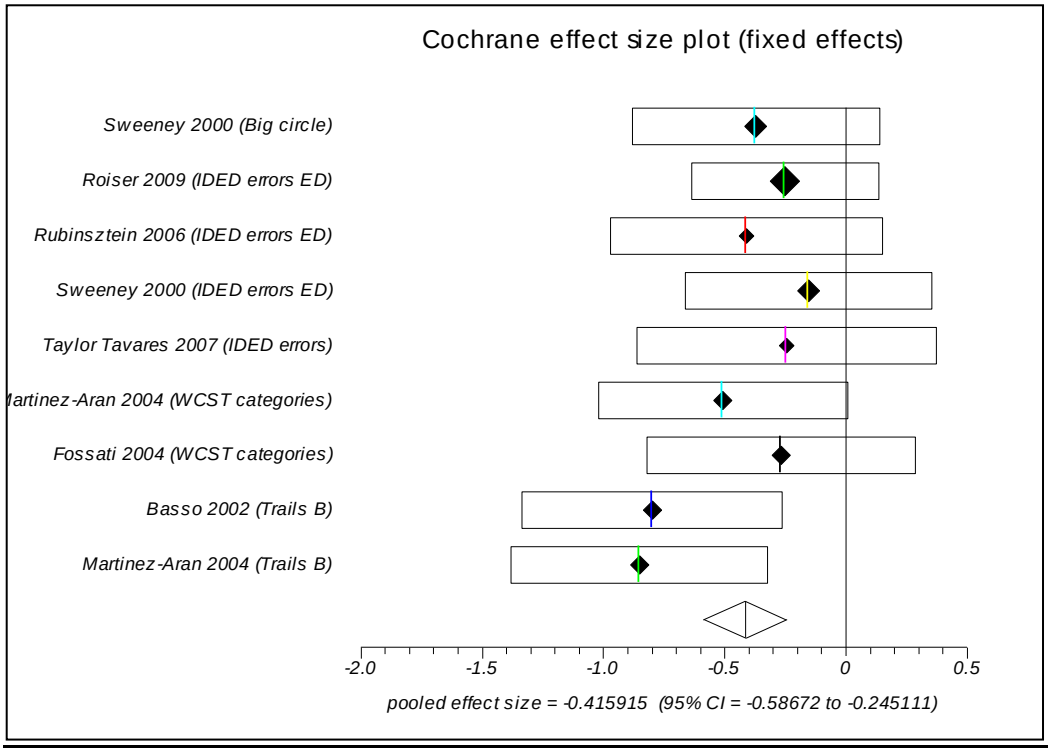
Effect size plot for RVIP latency



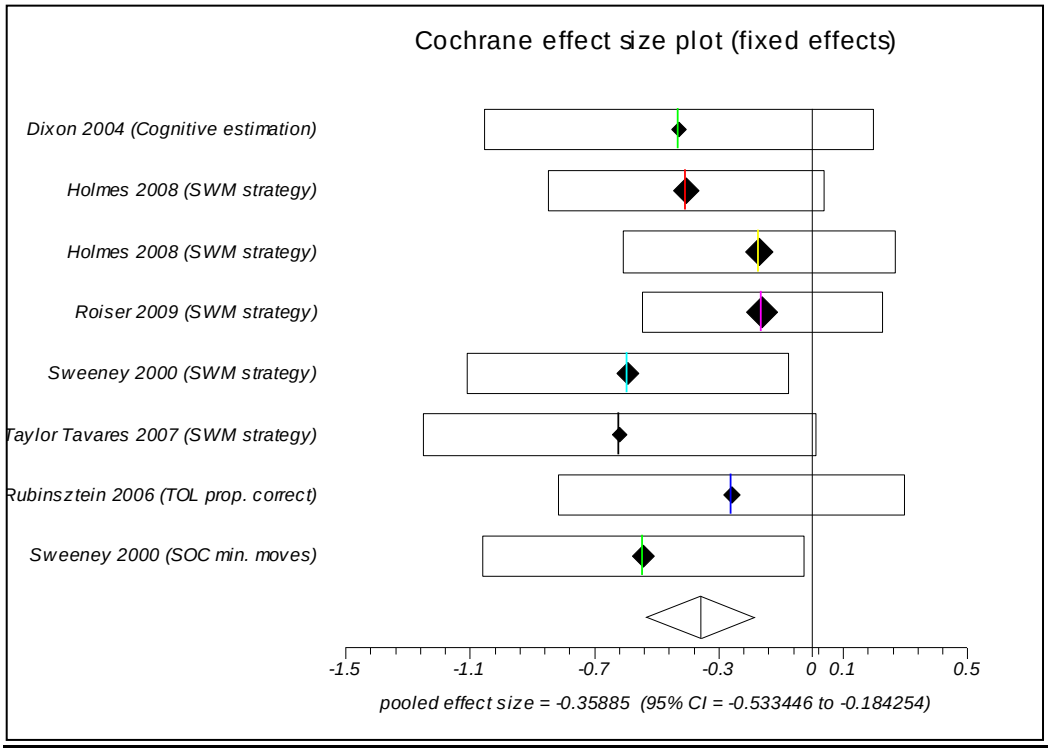
Effect size plot for executive functioning (working memory monitoring)



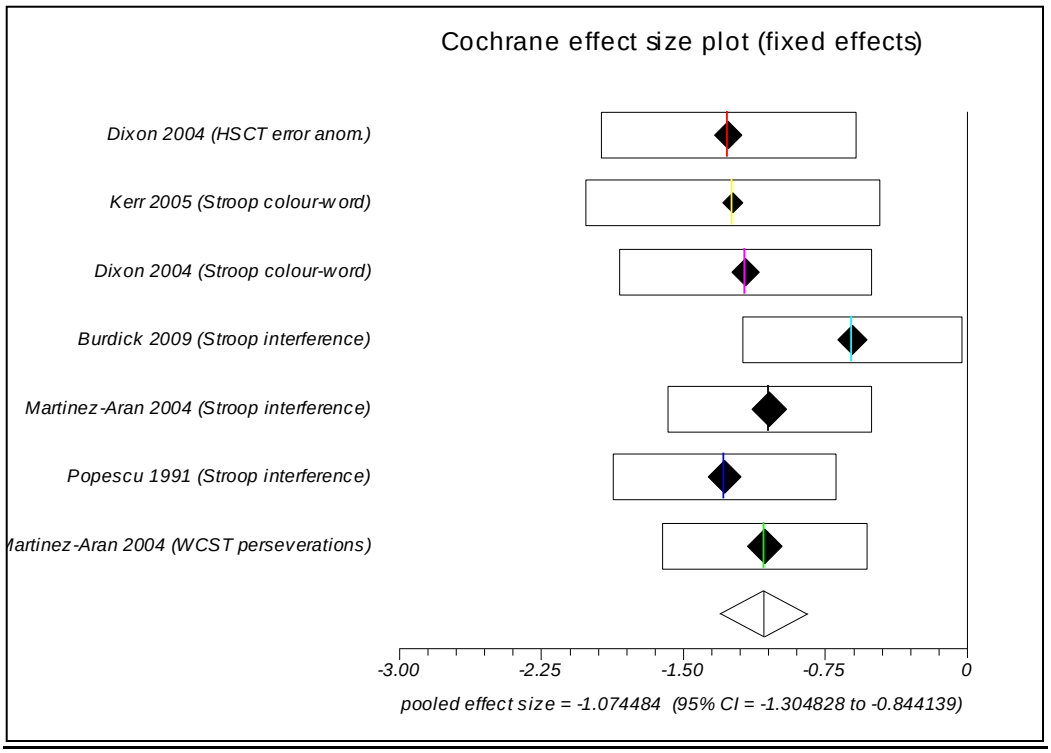
Effect size plot for executive functioning (Set shifting/ rule formation and reversal)



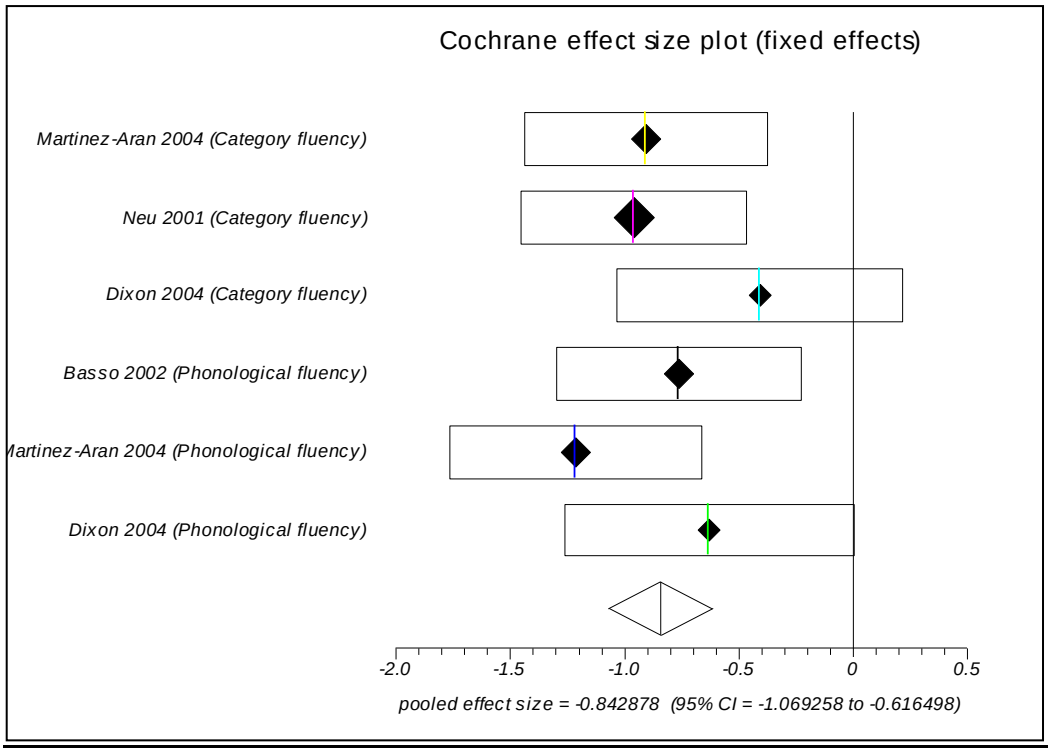
Effect size plot for executive functioning (Planning, reasoning and strategy)



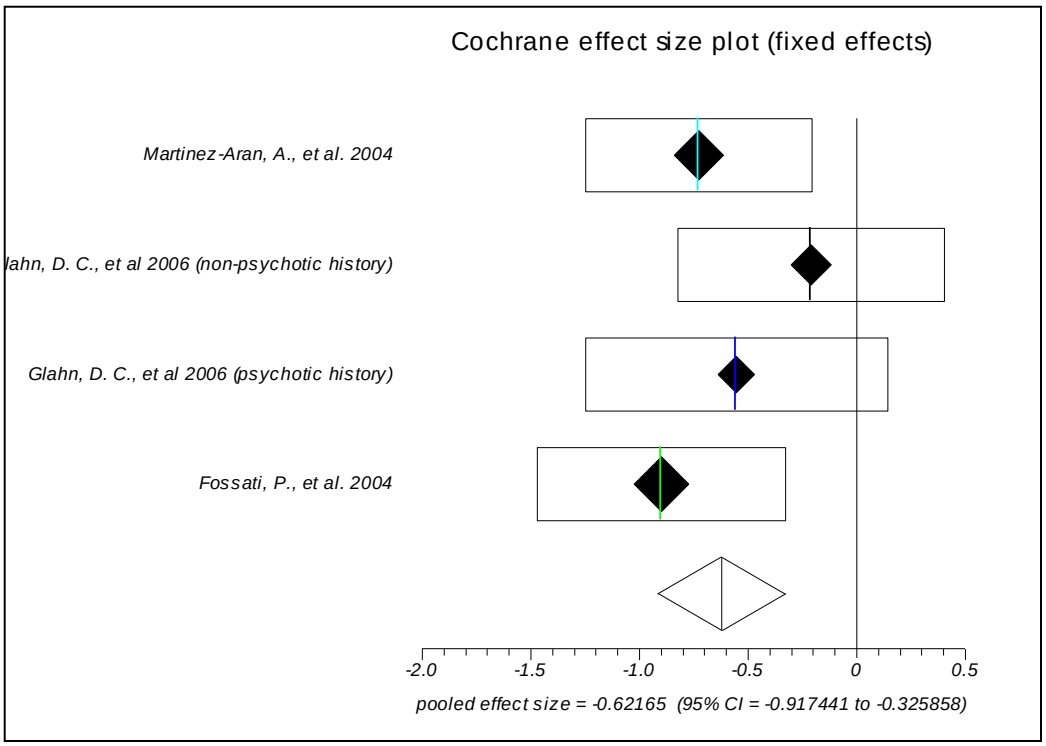
Effect size plot for executive functioning (Inhibition)



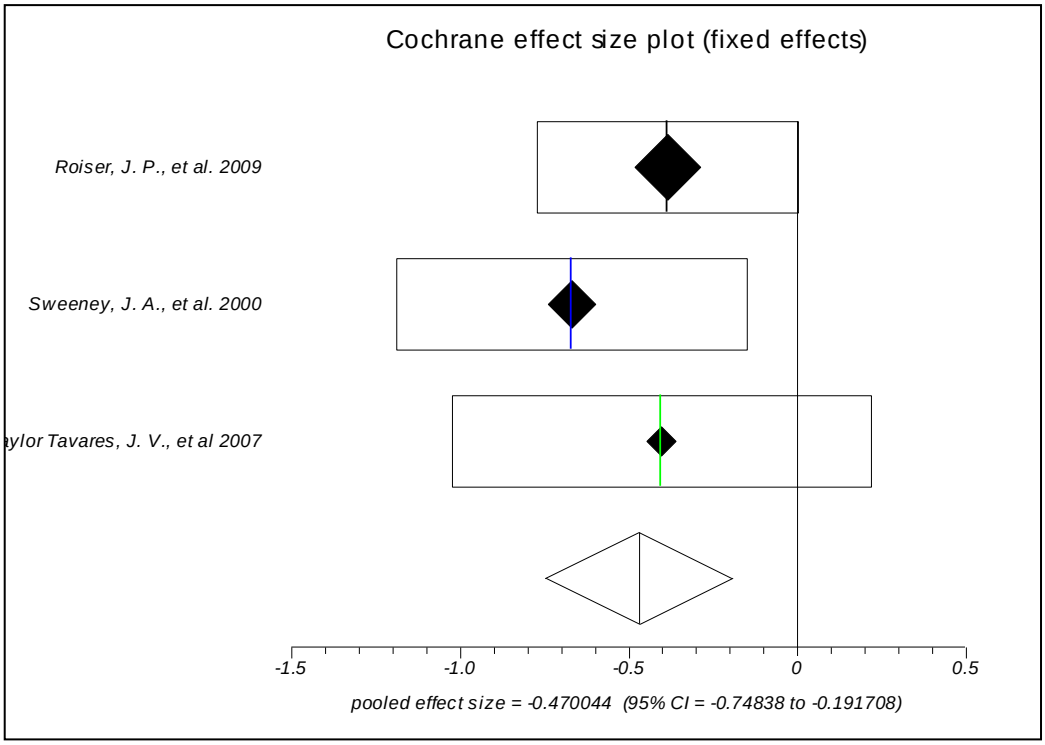
Effect size plot for executive functioning (Verbal fluency)



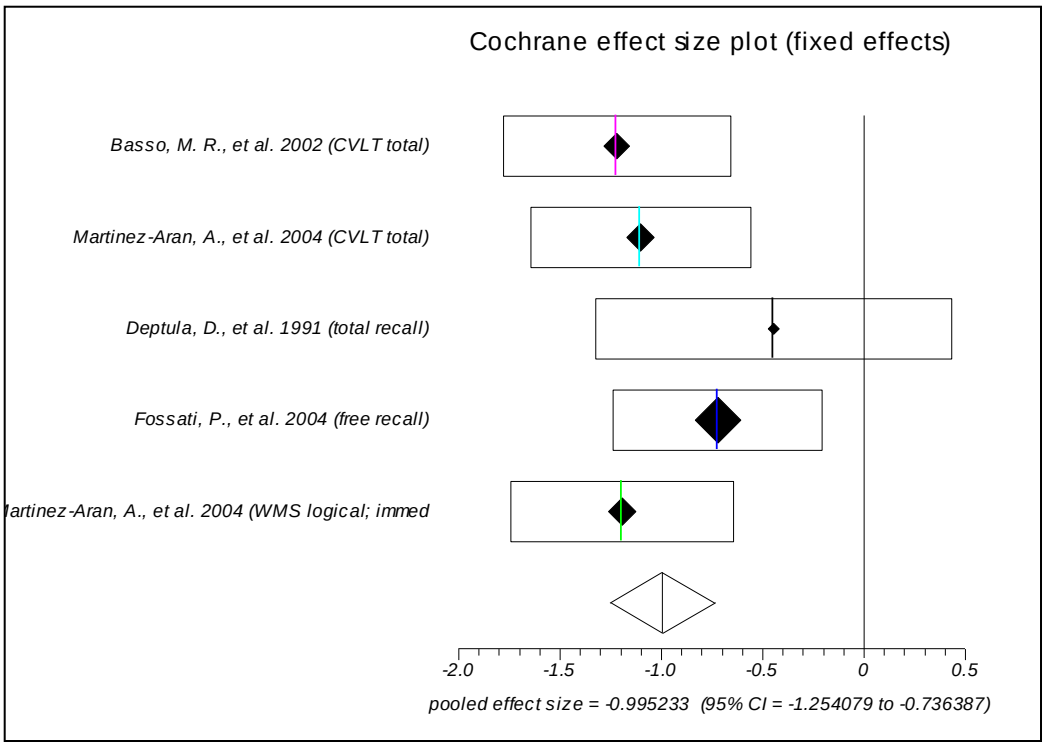
Effect size plot for digit span forwards



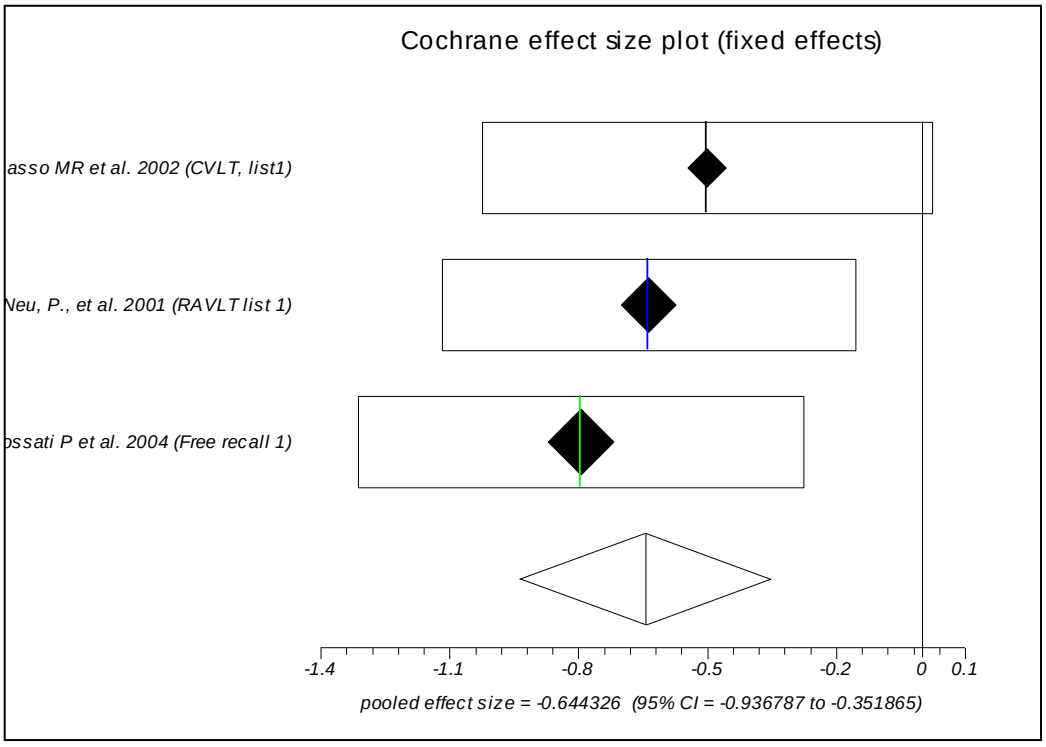
Effect size plot for spatial span forwards



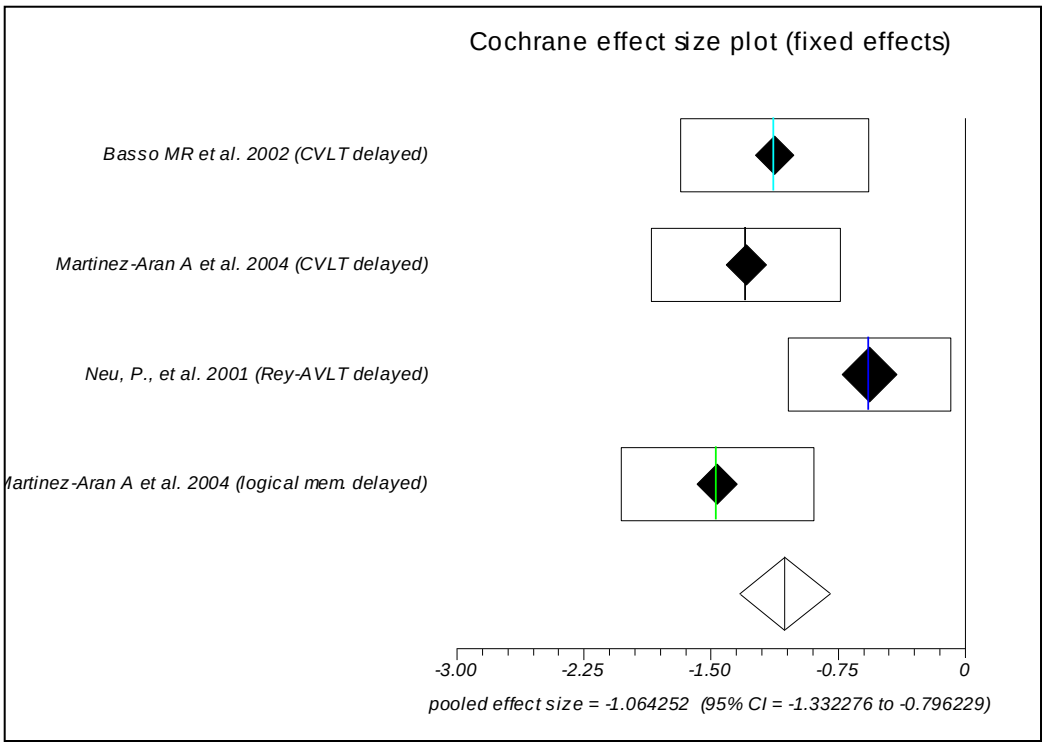
Effect size plot for total immediate free-recall



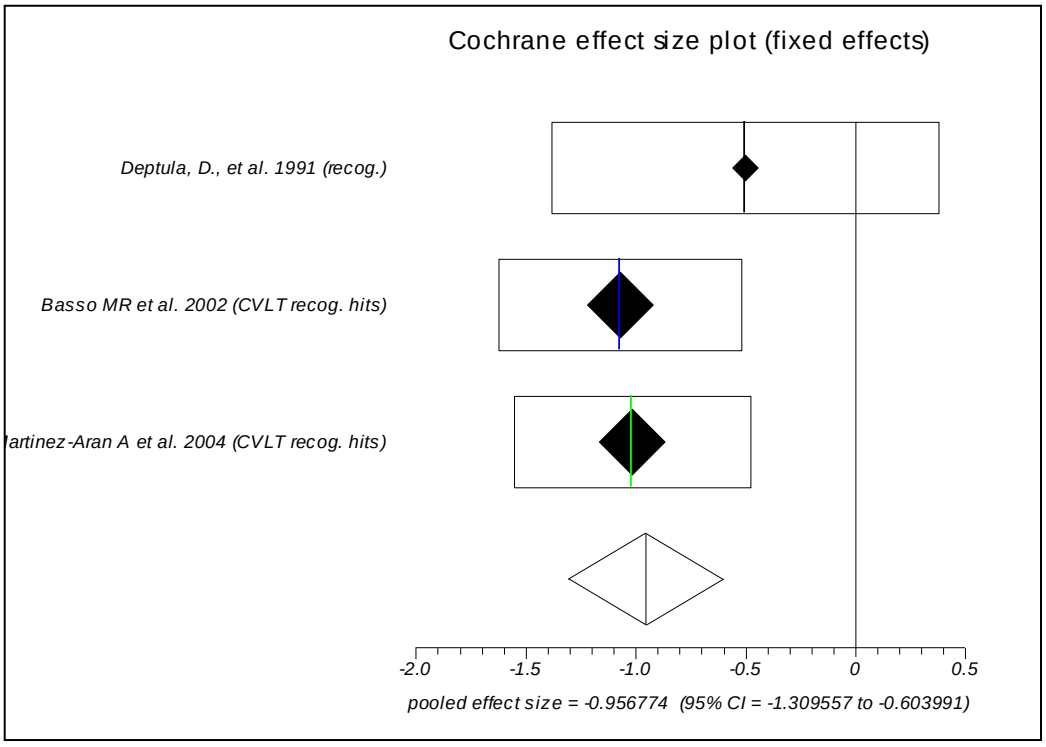
Effect size plot for initial immediate free-recall



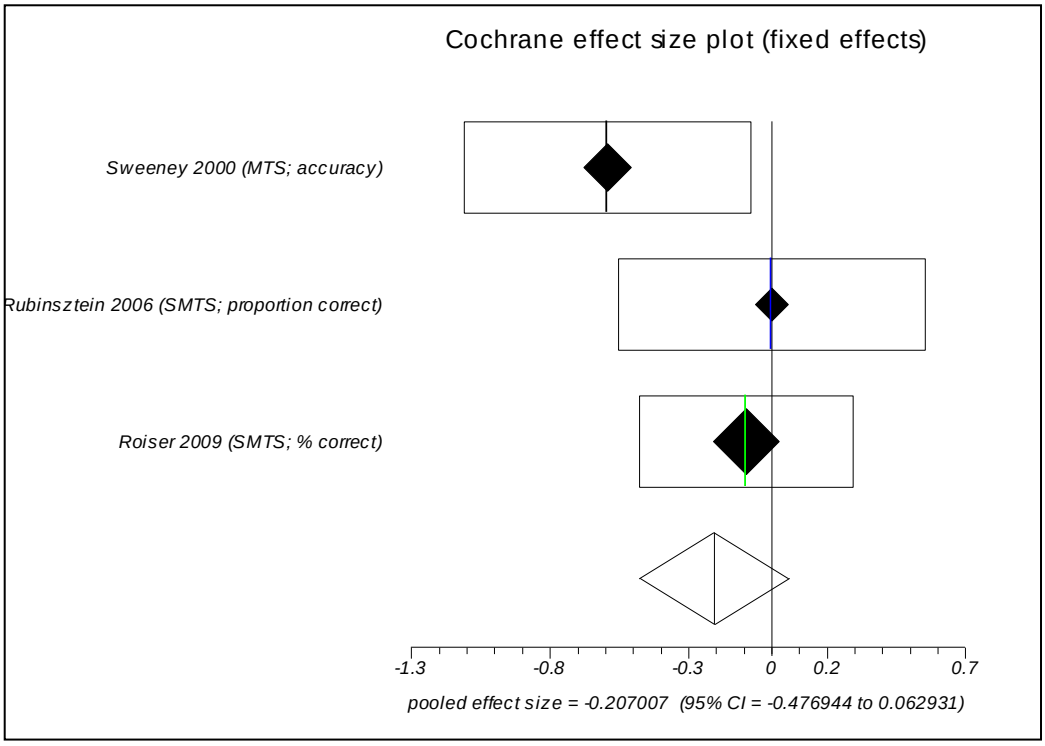
Effect size plot for delayed free-recall



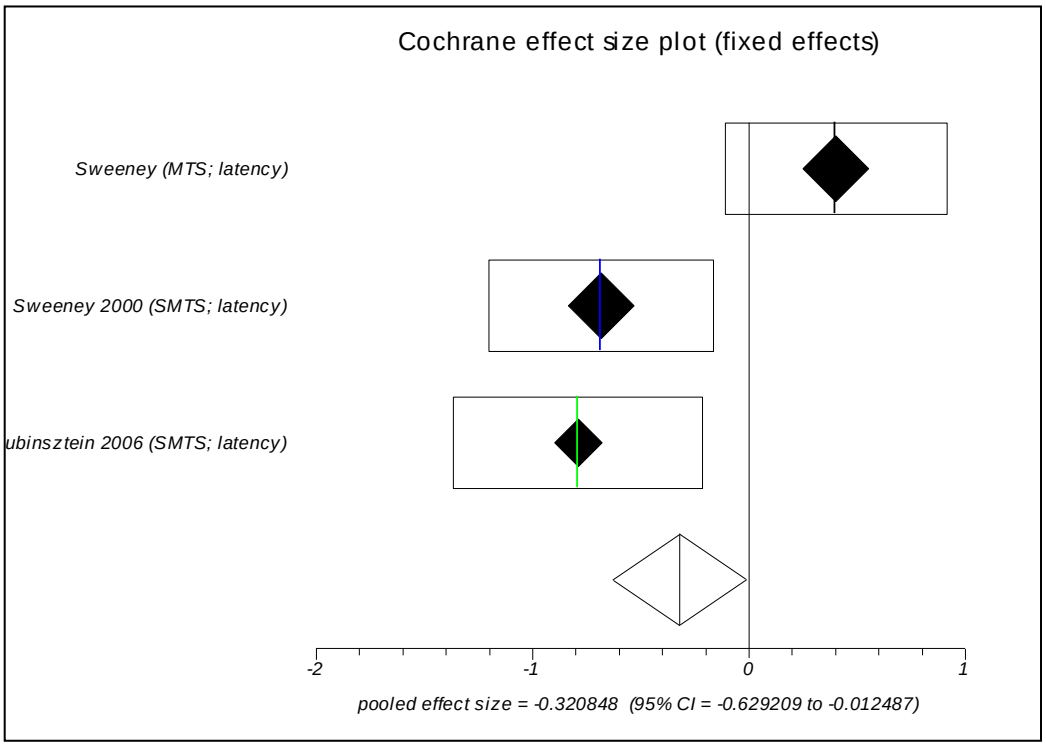
Effect size plot for delayed recognition



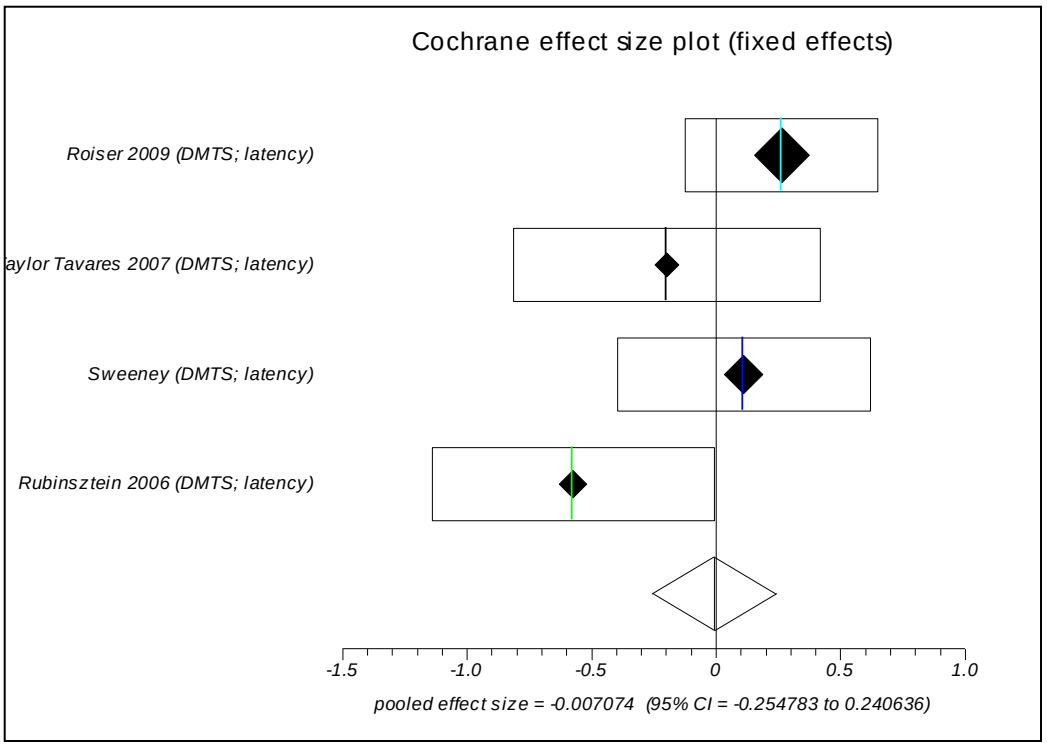
Effect size plot for simultaneous match-to-sample (accuracy)



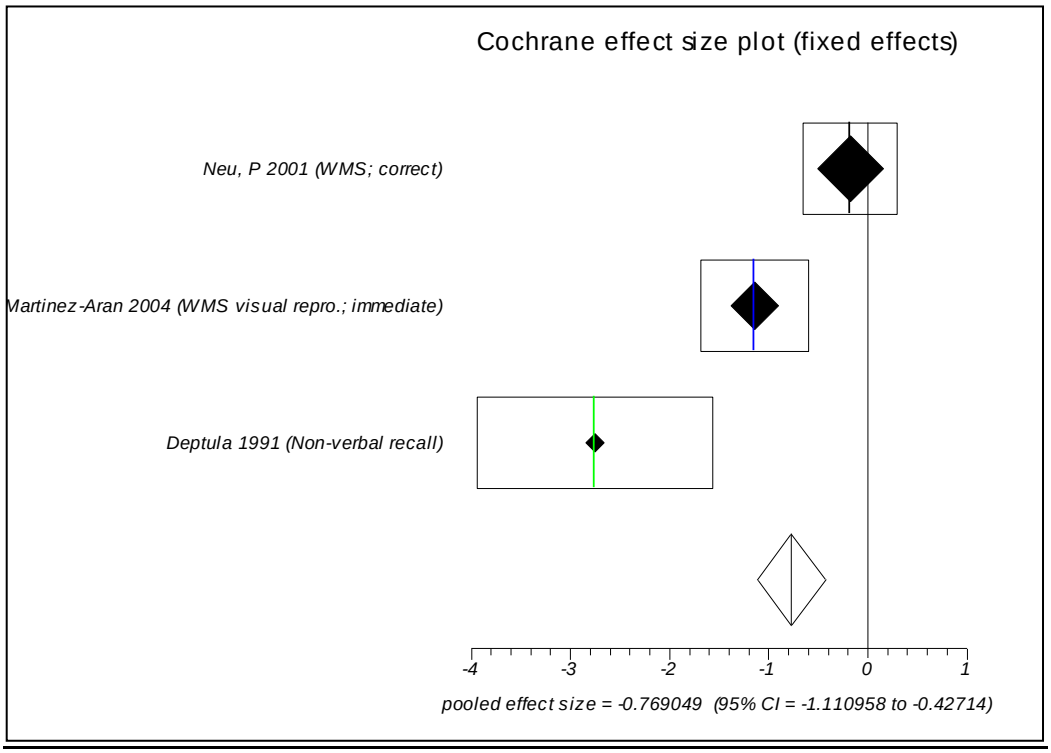
Effect size plot for simultaneous match-to-sample (latency)



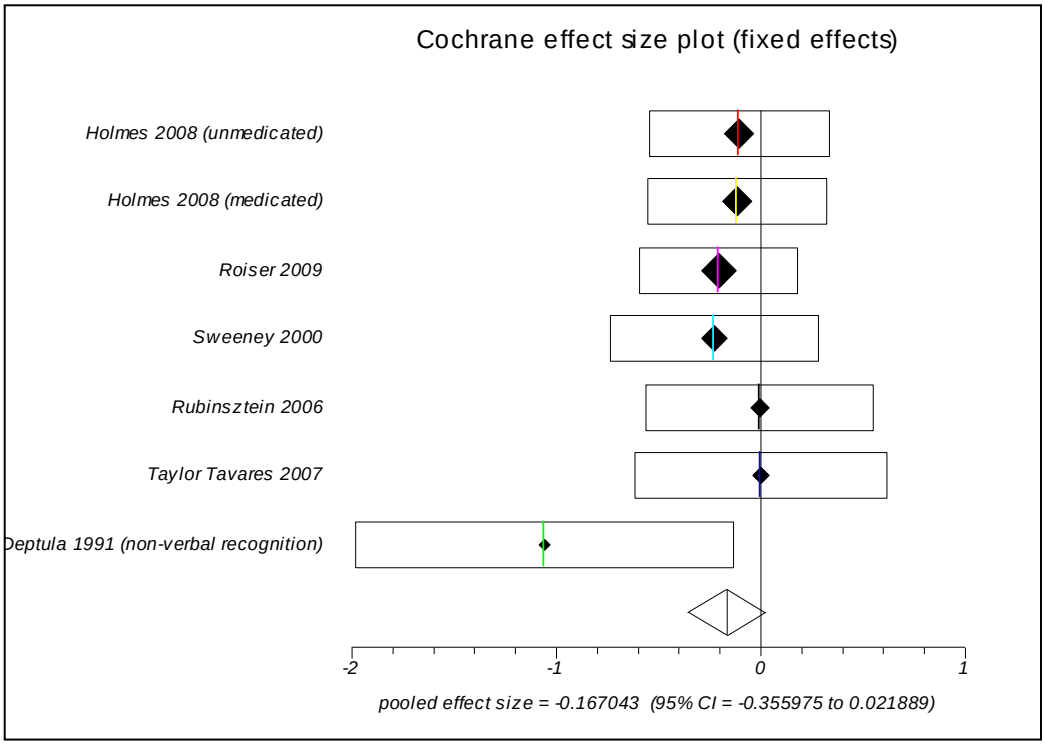
Effect size plot for delayed match-to-sample (latency)



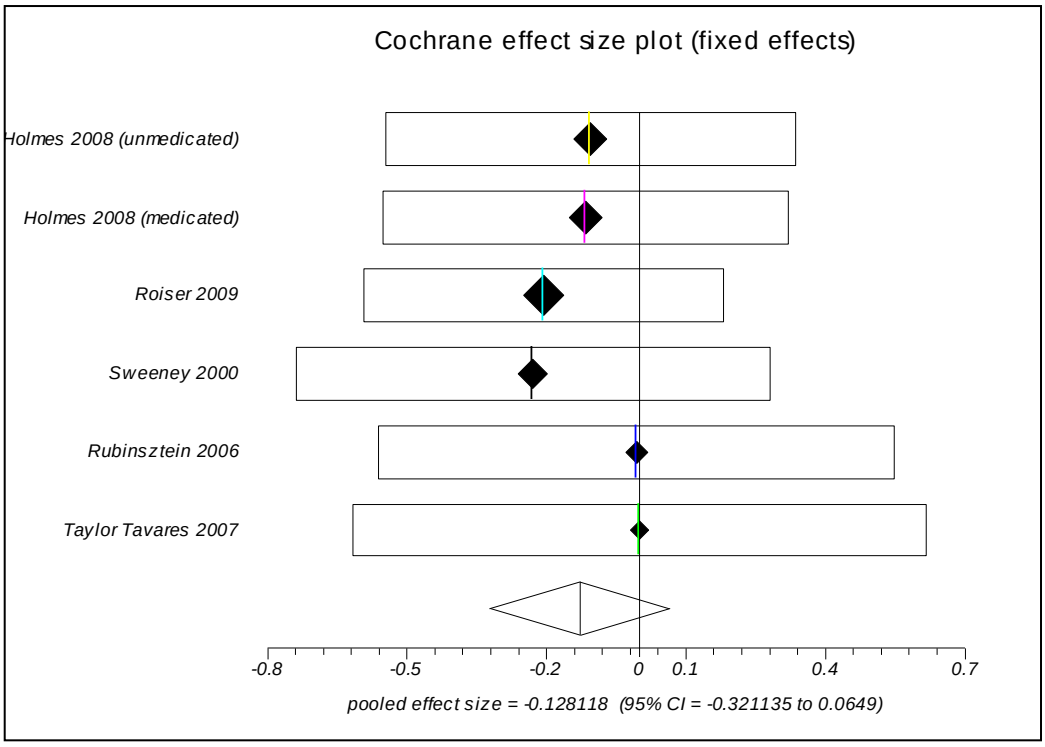
Effect size plot for visual memory (immediate recall)



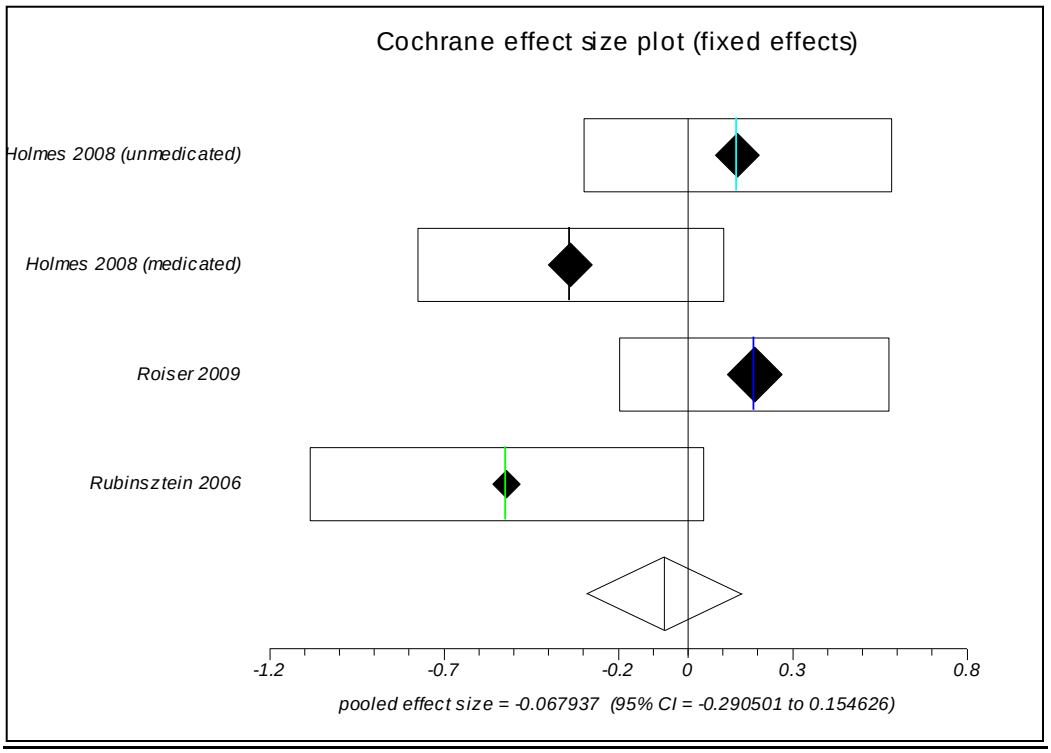
Effect size plot for visual memory (recognition)



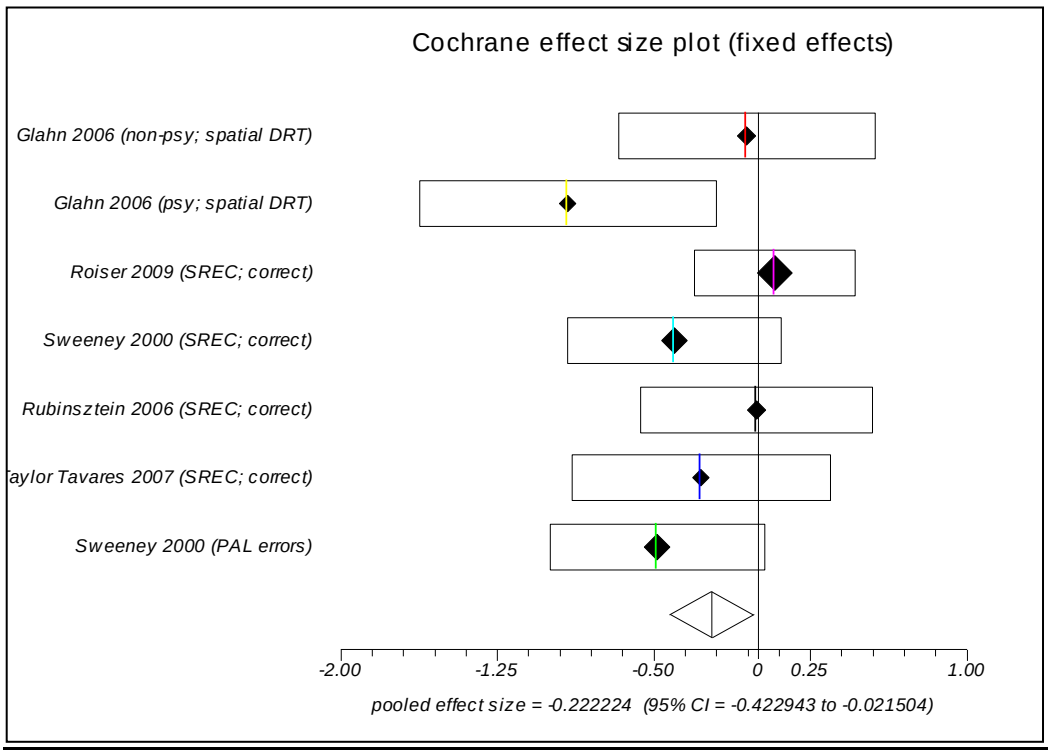
Effect size plot for visual memory (CANTAB PREC accuracy)



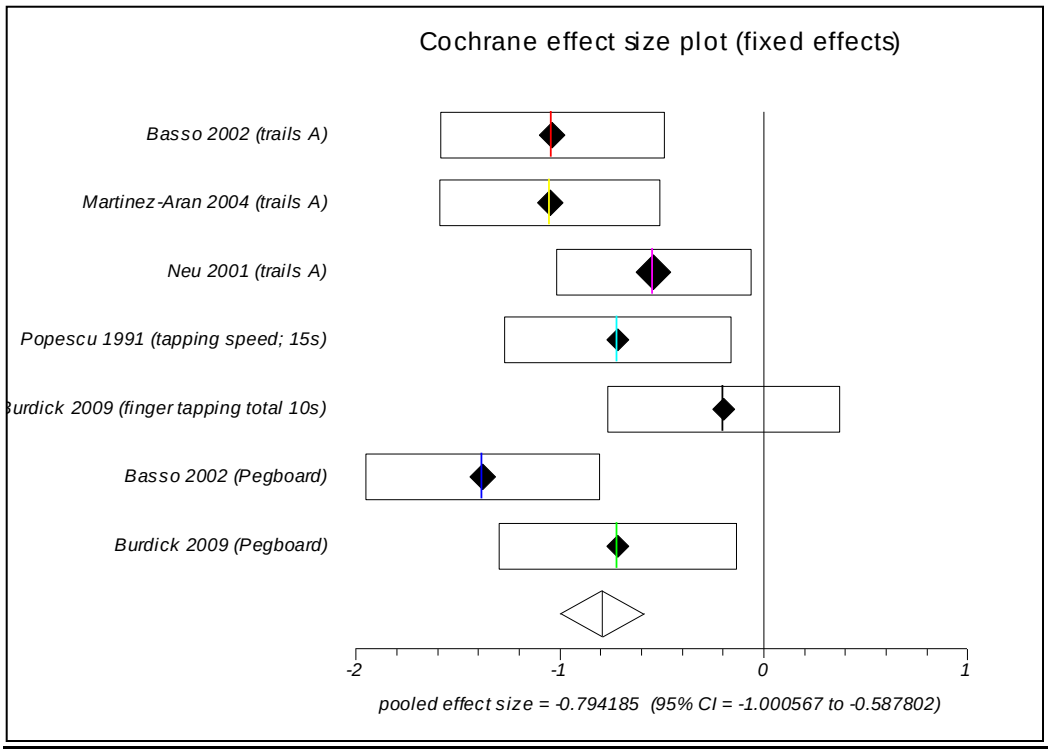
Effect size plot for visual memory (CANTAB PREC latency)



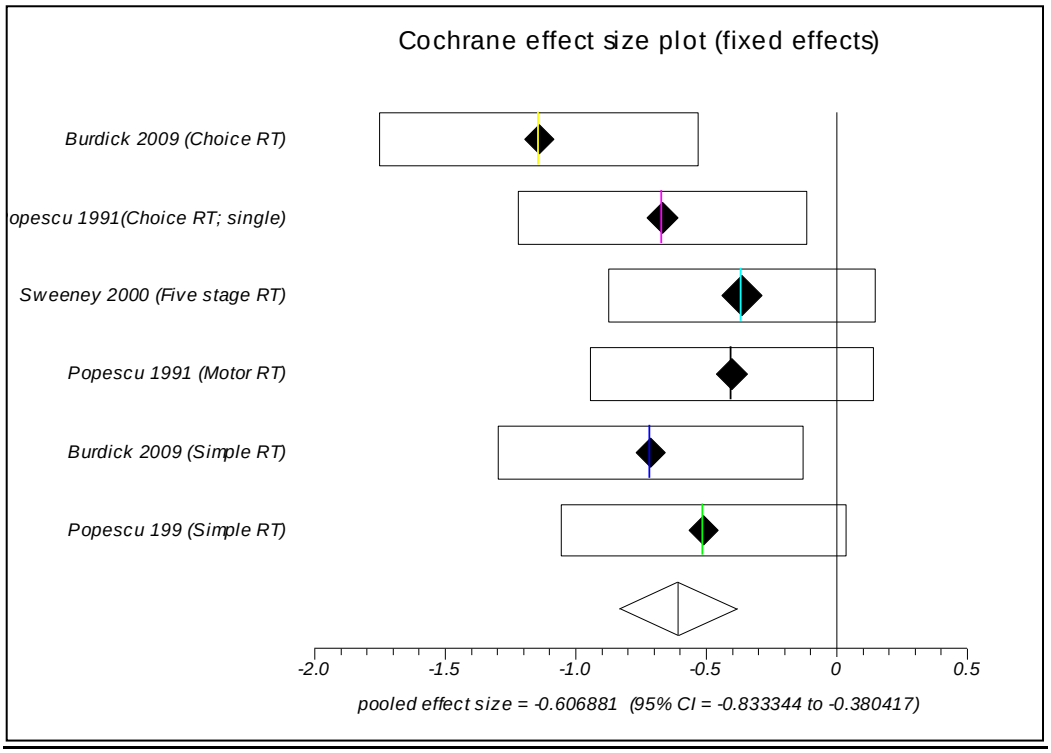
Effect size plot for spatial memory (accuracy)



Effect size plot for psychomotor tests (motor skills or dexterity)



Effect size plot for psychomotor tests (reaction time)



9.2 SAS matching of patients and controls (Chapter 2)

Data listing for matched Cases and Controls

	UNID	UNID	Control	Distance	Age	Sex	NART	Age	Age	Sex	Sex	NART	NART
Obs	Case	Control	id	D_IJ	Abs. Diff.	Abs. Diff.	Abs. Diff.	Case	Control	Case	Control	Case	Control
1	128	91	1	9	1	0	7	45	44	1	1	110	117
2	129	29	1	12	5	0	2	54	49	1	1	120	118
3	130	100	1	5	2	0	1	57	55	1	1	106	105
4	131	2	1	11	4	0	3	57	53	1	1	124	121
5	132	86	1	11.813	5	0	2	56	51	1	1	111	113
6	133	28	1	0.813	0	0	1	38	38	1	1	111	112
7	134	55	1	12	4	0	4	61	57	1	1	126	122
8	135	35	1	4	2	0	0	33	35	1	1	106	106
9	136	36	1	15	2	0	11	63	61	2	2	100	89
10	137	41	1	5	2	0	1	35	33	1	1	112	111
11	138	88	1	18.188	5	0	8	58	53	1	1	111	103
12	139	98	1	4	0	0	4	42	42	1	1	108	112
13	140	87	1	4	2	0	0	49	47	1	1	110	110
14	141	80	1	1.188	0	0	1	41	41	2	2	111	110
15	142	44	1	5	2	0	1	26	28	1	1	100	101
16	143	58	1	7	2	0	3	62	60	1	1	113	110
17	144	61	1	5	0	0	5	50	50	1	1	121	126
18	145	73	1	5	2	0	1	59	57	1	1	107	106
19	146	3	1	3	1	0	1	41	40	1	1	108	107
20	147	95	1	7	2	0	3	45	47	1	1	108	105
				========									
				145									

match macro: case=CASES control=CONTROLS idca=UNID idco=UNID mvars=AGE SEX NART wts=2 2 1 dmaxk=5 0 12 dmax= ncontls=1 method=optimal seedca= seedco= out=MTCH outnmca=__NMCA outnmco=__NMCO

The SAS System 16:51 Friday, September 5, 2008 8

Obs Variable	Label	N	Mean	Sum	Minimum	Maximum
DIJ	DISTANCE/D_IJ	20	7.25	145	0.8125	18.1875
DIF1	AGE/ABS. DIFF	20	2.15	43	0	5
DIF2	SEX/ABS. DIFF	20	0	0	0	0
DIF3	NART/ABS. DIFF	20	2.95	59	0	11
CA1	AGE/CASE	20	48.6	972	26	63
CA2	SEX/CASE	20	1.1	22	0	1
CA3	NART/CASE	20	111.1875	2223.75	100	126
CO1	AGE/CONTROL	20	47.05	941	28	61
CO2	SEX/CONTROL	20	1.1	22	1	2
CO3	NART/CONTROL	20	110.2	2204	89	126

9.3 Hamilton Rating Scale for Depression (HAM-D21)

To rate the severity of depression in patients who are already diagnosed as depressed, administer this questionnaire. The higher the score, the more severe the depression.

For each item, write the correct number on the line next to the item. (Only one response per item)

1. DEPRESSED MOOD (Sadness, hopeless, helpless, worthless)
0= Absent
1= These feeling states indicated only on questioning
2= These feeling states spontaneously reported verbally
3= Communicates feeling states non-verbally—i.e., through facial expression, posture, voice, and tendency to weep
4= Patient reports VIRTUALLY ONLY these feeling states in his spontaneous verbal and nonverbal communication

2. FEELINGS OF GUILT
0= Absent
1= Self reproach, feels he has let people down
2= Ideas of guilt or rumination over past errors or sinful deeds
3= Present illness is a punishment. Delusions of guilt
4= Hears accusatory or denunciatory voices and/or experiences threatening visual hallucinations

3. SUICIDE
0= Absent
1= Feels life is not worth living
2= Wishes he were dead or any thoughts of possible death to self
3= Suicidal ideas or gesture
4= Attempts at suicide (any serious attempt rates 4)

4. INSOMNIA EARLY
0= No difficulty falling asleep
1= Complains of occasional difficulty falling asleep—i.e., more than 1/2 hour
2= Complains of nightly difficulty falling asleep

5. INSOMNIA MIDDLE
0= No difficulty
1= Patient complains of being restless and disturbed during the night
2= Waking during the night—any getting out of bed rates 2 (except for purposes of voiding)

6. INSOMNIA LATE
0= No difficulty
1= Waking in early hours of the morning but goes back to sleep
2= Unable to fall asleep again if he gets out of bed

7. WORK AND ACTIVITIES
0= No difficulty
1= Thoughts and feelings of incapacity, fatigue or weakness related to activities; work or hobbies
2= Loss of interest in activity; hobbies or work—either directly reported by patient, or indirect in listlessness, indecision and vacillation (feels he has to push self to work or activities)
3= Decrease in actual time spent in activities or decrease in productivity
4= Stopped working because of present illness

8. RETARDATION: PSYCHOMOTOR (Slowness of thought and speech; impaired ability to concentrate; decreased motor activity)
0= Normal speech and thought
1= Slight retardation at interview
2= Obvious retardation at interview
3= Interview difficult
4= Complete stupor

9. AGITATION
0= None
1= Fidgetiness
2= Playing with hands, hair, etc.
3= Moving about, can't sit still
4= Hand wringing, nail biting, hair-pulling, biting of lips

10. ANXIETY (PSYCHOLOGICAL)
0= No difficulty
1= Subjective tension and irritability
2= Worrying about minor matters
3= Apprehensive attitude apparent in face or speech
4= Fears expressed without questioning

11. ANXIETY SOMATIC: Physiological concomitants of anxiety, (i.e., effects of autonomic overactivity, "butterflies," indigestion, stomach cramps, belching, diarrhoea, palpitations, hyperventilation, paresthesia, sweating, flushing, tremor, headache, urinary frequency).
Avoid asking about possible medication side effects (i.e., dry mouth, constipation)
0= Absent
1= Mild
2= Moderate
3= Severe
4= Incapacitating

12. SOMATIC SYMPTOMS (GASTROINTESTINAL)
0= None
1= Loss of appetite but eating without encouragement from others. Food intake about normal
2= Difficulty eating without urging from others. Marked reduction of appetite and food intake

13. SOMATIC SYMPTOMS GENERAL
0= None
1= Heaviness in limbs, back or head. Backaches, headache, muscle aches. Loss of energy and fatigability
2= Any clear-cut symptom rates 2

14. GENITAL SYMPTOMS (Symptoms such as: loss of libido; impaired sexual performance; menstrual disturbances)
0= Absent
1= Mild
2= Severe

15. HYPOCHONDRIASIS
0= Not present
1= Self-absorption (bodily)
2= Preoccupation with health
3= Frequent complaints, requests for help, etc.
4= Hypochondriacal delusions

16. LOSS OF WEIGHT
A. When rating by history:
0= No weight loss
1= Probably weight loss associated with present illness
2= Definite (according to patient) weight loss
3= Not assessed

17. INSIGHT
0= Acknowledges being depressed and ill
1= Acknowledges illness but attributes cause to bad food, climate, overwork, virus, need for rest, etc.
2= Denies being ill at all

SCORE FOR HAM-D 17: _____

Additional items for HAMD-21

18. DIURNAL VARIATION

A. Note whether symptoms are worse in morning or evening. If NO diurnal variation, mark none

0= No variation
1= Worse in A.M.
2= Worse in P.M.

B. When present, mark the severity of the variation. Mark “None” if NO variation

0= None
1= Mild
2= Severe

19. DEPERSONALIZATION AND DEREALIZATION (Such as: Feelings of unreality; Nihilistic ideas)

0= Absent
1= Mild
2= Moderate
3= Severe
4= Incapacitating

20. PARANOID SYMPTOMS

0= None
1= Suspicious
2= Ideas of reference
3= Delusions of reference and persecution

21. OBSESSIONAL AND COMPULSIVE SYMPTOMS

0= Absent
1= Mild
2= Severe

Total Score (HAM-D 21) _____

9.4 Montgomery and Åsberg (MADRS) Depression Rating Scale

The rating should be based on a clinical interview moving from broadly phrased questions about symptoms to more detailed ones which allow a precise rating of severity. The rater must decide whether the rating lies on the defined scale steps (0, 2, 4, 6) or between them (1, 3, 5).
It is important to remember that it is only on rare occasions that a depressed patient is encountered who cannot be rated on the items in the scale. If definite answers cannot be elicited from the patient all relevant clues as well as information from other sources should be used as a basis for the rating in line with customary clinical practice.
The scale may be used for any time interval between ratings, be it weekly or otherwise but this must be recorded.

Item List

1. Apparent sadness
2. Reported sadness
3. Inner tension
4. Reduced sleep
5. Reduced appetite
6. Concentration difficulties
7. Lassitude
8. Inability to feel
9. Pessimistic thoughts
10. Suicidal thoughts

1. Apparent Sadness

Representing despondency, gloom and despair, (more than just ordinary transient low spirits) reflected in speech, facial expression, and posture. Rate by depth and inability to brighten up.

0 No sadness.
1
2 Looks dispirited but does brighten up without difficulty.
3
4 Appears sad and unhappy most of the time.
5
6 Looks miserable all the time. Extremely despondent.

2. Reported sadness

Representing reports of depressed mood, regardless of whether it is reflected in appearance or not. Includes low spirits, despondency or the feeling of being beyond help and without hope.

Rate according to intensity, duration and the extent to which the mood is reported to be influenced by events.

0 Occasional sadness in keeping with the circumstances.

1

2 Sad or low but brightens up without difficulty.

3

4 Pervasive feelings of sadness or gloominess. The mood is still influenced by external circumstances.

5

6 Continuous or unvarying sadness, misery or despondency.

3. Inner tension

Representing feelings of ill-defined discomfort, edginess, inner turmoil, mental tension mounting to either panic, dread or anguish.

Rate according to intensity, frequency, duration and the extent of reassurance called for.

0 'Placid. Only fleeting inner tension.

1

2 Occasional feelings of edginess and ill¬defined discomfort.

3

4 Continuous feelings of inner tension or intermittent panic which the patient can only master with some difficulty.

5

6 Unrelenting dread or anguish. Overwhelming panic.

4. Reduced sleep

Representing the experience of reduced duration or depth of sleep compared to the subject's own normal pattern when well.

0 Sleeps as usual.

1

2 Slight difficulty dropping off to sleep or slightly reduced, light or fitful sleep.

3

4 Sleep reduced or broken by at least two hours.

5

6. Less than two or three hours sleep.

5. Reduced appetite

Representing the feeling of a loss of appetite compared with when well. Rate by loss of desire for food or the need to force oneself to eat.

0 Normal or increased appetite.
1
2 Slightly reduced appetite.
3
4 No appetite. Food is tasteless.
5
6 Needs persuasion to eat at all.

6. Concentration difficulties

Representing difficulties in collecting one's thoughts mounting to incapacitating lack of concentration. Rate according to intensity, frequency, and degree of incapacity produced.

0 No difficulties in concentrating.
1
2 Occasional difficulties in collecting one's thoughts.
3
4 Difficulties in concentrating and sustaining thought which reduces ability to read or hold a conversation.
5
6 Unable to read or converse without great difficulty.

7. Lassitude

Representing a difficulty getting started or slowness initiating and performing everyday activities.

0 Hardly any difficulty in getting started. No sluggishness.
1
2 Difficulties in starting activities.
3
4 Difficulties in starting simple routine activities which are carried out with effort.
5
6 Complete lassitude. Unable to do anything without help.

8. Inability to feel

Representing the subjective experience of reduced interest in the surroundings, or activities that normally give pleasure. The ability to react with adequate emotion to circumstances or people is reduced.

0 Normal interest in the surroundings and in other people.
1
2 Reduced ability to enjoy usual interests. 3
4 Loss of interest in the surroundings. Loss of feelings for friends and acquaintances.
5
6 The experience of being emotionally para¬lysed, inability to feel anger, grief or pleasure and a complete or even painful failure to feel for close relatives and friends.

9. Pessimistic thoughts

Representing thoughts of guilt, inferiority, self¬-reproach, sinfulness, remorse and ruin.

0 No pessimistic thoughts.
1
2 Fluctuating ideas of failure, self-reproach or self depreciation.
3
4 Persistent self-accusations, or definite but still rational ideas of guilt or sin. Increasingly pessimistic about the future.
5
6 Delusions of ruin, remorse or unredeemable sin. Self-accusations which are absurd and unshakable.

10. Suicidal thoughts

Representing the feeling that life is not worth living, that a natural death would be welcome, suicidal thoughts, and preparations for suicide.
Suicidal attempts should not in themselves influence the rating.

0 Enjoys life or takes it as it comes.
1
2 Weary of life. Only fleeting suicidal thoughts.
3
4 Probably better off dead. Suicidal thoughts are common, and suicide is considered as a possible solution, but without specific plans or intention.
5
6 Explicit plans for suicide when there is an opportunity. Active preparations for suicide.

9.5 Beck Depression Inventory (BDI)

On this questionnaire are groups of statements. Please read each group of statements carefully, then pick out the one statement in each group which best describes the way you have been feeling over the **PAST WEEK**. Circle the number beside the statement you picked. If several statements in the group seem to apply equally well then circle each one. **Be sure to read all the statements in each group before making your choice.**

1)
0 I do not feel sad.
1 I feel sad.
2 I am sad all the time and can't snap out of it.
3 I am so sad or unhappy that I can't stand it.

2)
0 I am not particularly discouraged about the future.
1 I feel discouraged about the future.
2 I feel I have nothing to look forward to.
3 I feel the future is hopeless and that things cannot improve.

3)
0 I do not feel like a failure.
1 I feel I have failed more than the average person.
2 As I look back on my life, all I can see is a lot of failures.
3 I feel I am a complete failure as a person.

4)
0 I get as much satisfaction out of things as I used to.
1 I don't enjoy things the way I used to.
2 I don't get real satisfaction out of anything anymore.
3 I am dissatisfied or bored with everything.

5)
0 I don't feel particularly guilty.
1 I feel guilty a good part of the time.
2 I feel quite guilty most of the time.
3 I feel guilty all of the time.

6)
0 I don't feel I am being punished.
1 I feel I may be punished.
2 I expect to be punished.
3 I feel I am being punished.

7)
0 I don't feel disappointed in myself.
1 I am disappointed in myself.
2 I am disgusted with myself.
3 I hate myself.

8)
0 I don't feel I am worse than anybody else.
1 I am critical of myself for my weaknesses or mistakes.
2 I blame myself all the time for my faults.
3 I blame myself for everything bad that happens.

9)
0 I don't have any thoughts of killing myself.
1 I have thoughts of killing myself, but I would not carry them out.
2 I would like to kill myself.
3 I would kill myself if I had the chance.

10)
0 I don't cry any more than usual.
1 I cry more now than I used to.
2 I cry all the time now.
3 I used to be able to cry, but now I can't even cry even though I want to.

11)
0 I am no more irritated by things than I ever am.
1 I get annoyed or irritated more easily than I used to.
2 I feel irritated all the time now.
3 I don't get irritated at all by things that used to irritate me.

12)
0 I have not lost interest in other people.
1 I am less interested in other people than I used to be.
2 I have lost most of my interest in other people.
3 I have lost all of my interest in other people.

13)
0 I make decisions about as well as I ever could.
1 I put off making decisions more than I used to.
2 I have greater difficulty in making decisions than before.
3 I can't make decisions at all anymore.

14)
0 I don't feel that I look any worse than I used to.
1 I am worried that I am looking old or unattractive.
2 I feel that there are permanent changes in my appearance that make me look unattractive.
3 I believe that I look ugly.

15)
0 I can work about as well as before.
1 It takes an extra effort to get started at doing something.
2 I have to push myself very hard to do anything.
3 I can't do any work at all.

16)
0 I can sleep as well as usual.
1 I don't sleep as well as I used to.
2 I wake up 1-2 hours earlier than usual and find it hard to get back to sleep.
3 I wake up several hours earlier than I used to and cannot get back to sleep.

17)
0 I don't get tired more than usual.
1 I get tired more easily than I used to.
2 I get tired from doing almost anything.
3 I am too tired to do anything.

18)
0 My appetite is no worse than usual.
1 My appetite is not as good as it used to be.
2 My appetite is much worse now.
3 I have no appetite at all anymore.

19)
0 I haven't lost any weight recently.
1 I have lost more than five pounds.
2 I have lost more than ten pounds.
3 I have lost more than fifteen pounds.
I am purposefully trying to lose weight by eating less.

20)
0 I am no more worried about my health than usual.
1 I am worried about physical problems such as aches or pains, or upset stomach, or constipation.
2 I am very worried about physical problems and it's hard to think of much else.
3 I am so worried about my physical problems that I cannot think about anything else.

21)
0 I have not noticed any recent change in my interest in sex.
1 I am less interested in sex than I used to be.
2 I am much less interested in sex now.
3 I have lost interest in sex completely.

9.6 Correlations between HPA axis measures and composites

Whole group	Salivary C_D ratio 8am	Salivary C_D ratio 8pm	Waking CORT AUC	DST CORT change
Salivary C_D ratio 8pm	0.199			
Waking CORT AUC	**0.299***	-0.146		
DST CORT change (2-1)	0.068	0.072	-0.269	
Optimised model_TOTAL_c1 (VS)	-0.049	0.000	-0.002	0.063
Optimised model_TOTAL_c2 (verbal memory)	0.100	0.020	0.042	-0.181
Optimised model_TOTAL_c3 (verbal exec)	0.023	0.058	0.042	-0.106
Optimised SWM model_TOTAL_c1 (verbal memory)	0.100	0.020	0.042	-0.181
Optimised SWM model_TOTAL_c2 (VS)	-0.055	-0.137	-0.001	.082
Optimised SWM model_TOTAL_c3 (verbal exec)	0.023	0.058	0.042	-0.106

* $p<0.05$

By group	Group	Salivary C_D ratio 8am	Salivary C_D ratio 8pm	Waking CORT AUC	DST CORT change (2-1)
Salivary C_D ratio 8pm	Patient	0.169			
	Control	0.229			
Waking CORT AUC	Patient	0.202	**-0.488***		
	Control	**0.373***	0.083		
DST CORT change (2-1)	Patient	0.383	-0.150	-0.220	
	Control	0.007	0.216	-0.197	
Optimised model_control_c1 (verbal memory)	Patient	0.022	0.125	-0.056	-0.084
	Control	0.278	0.022	0.072	0.046
Optimised model_control_c2 (VS complex)	Patient	0.028	0.168	-0.111	-0.043
	Control	0.074	0.188	0.144	-0.017
Optimised model_control_c3 (VS immediate)	Patient	0.263	-0.174	-0.076	**0.420***
	Control	-0.279	-0.168	-0.049	0.112
Optimised model_control_c4 (verbal exec)	Patient	0.033	0.097	-0.201	0.168
	Control	0.271	0.257	0.219	-0.145
Optimised model_patient_c1 (VS)	Patient	0.092	-0.038	-0.112	0.103
	Control	-0.126	0.023	0.044	0.083
Optimised model_patient_c2 (verbal memory)	Patient	0.022	0.125	-0.056	-0.084
	Control	0.278	0.022	0.072	0.046
Optimised model_patient_c3 (digit srec)	Patient	-0.055	0.148	-0.290	-0.147
	Control	0.161	0.019	0.180	-0.053
Optimised SWM model_control_c1 (verbal memory)	Patient	0.022	0.125	-0.056	-0.084
	Control	0.278	0.022	0.072	0.046
Optimised SWM model_control_c2 (verbal exec)	Patient	0.011	0.025	-0.199	0.124
	Control	0.165	0.152	0.179	-0.101
Optimised SWM model_control_c3 (VS)	Patient	0.157	-0.197	-0.084	0.194
	Control	-0.255	-0.132	-0.012	0.160
Optimised SWM model_patient_c1 (VS)	Patient	0.102	-0.184	-0.098	0.176
	Control	-0.126	-0.018	0.021	0.178
Optimised SWM model_patient_c2 (verbal memory)	Patient	0.022	0.125	-0.056	-0.084
	Control	0.278	0.022	0.072	0.046
Optimised SWM model_patient_c3 (digit srec)	Patient	-0.055	0.148	-0.290	-0.147
	Control	0.161	0.019	0.180	-0.053

* p<0.05