Core cognitive impairments and their neural correlates in mood disorders

Bethany Ann Little

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

Translational and Clinical Research Institute

Faculty of Medical Sciences

Newcastle University

December 2022

ABSTRACT

Background: Mood disorders such as bipolar disorder (BD) and major depressive disorder (MDD) are associated with impairments in a wide range of cognitive functions compared to healthy controls (HCs). However, it is unclear whether the this reflects a general cognitive impairment, or whether specific primary impairments are responsible for wider dysfunction. Processing speed (PS), sustained attention (SA), and executive function (EF) may be particularly impaired in mood disorders and may affect wider cognition, creating a hierarchy of cognitive dysfunction. Mixed findings in the literature are partly due to heterogeneity of neuropsychological assessment across studies. Studies rarely account for relationships between cognitive functions or investigate the presence of a cognitive hierarchy. Cognitive impairments are also thought to relate to structural brain abnormalities in mood disorders, but whether specific functions are related to distinct brain areas is not clear and much of the research on brain-cognition associations is limited by univariate analysis.

Methods: A systematic review of *k*=103 studies was conducted to examine the presence and magnitude of PS and SA impairments in people with BD and MDD compared to HCs. Data were meta-analysed for each neuropsychological test score separately and subgroup analysis was performed across mood states, where possible. Hierarchical regression was used to examine the role of PS, SA, and EF in memory in euthymic BD (*n*=62), BD depression (*n*=43), and MDD (*n*=41) compared to matched HCs (*n*=142), controlling for age and premorbid IQ. Ex-Gaussian distributional parameters were obtained from continuous performance test reaction times to conceptualise SA. Network graphs were used to illustrate interrelationships between cognitive functions within each group. Canonical correlation analysis was used to assess multivariate associations between abnormal cortical thickness and core cognitive impairments in euthymic BD (*n*=56), controlling for age, sex, and premorbid IQ.

Results: The meta-analysis suggested that BD and MDD show impairments in PS and SA across most neuropsychological tests. Impairments were present in both symptomatic states and in euthymia in most cases, however, some outcome measures were not impaired in euthymia. PS and EF appeared to explain memory impairments in euthymic BD, whereas SA appeared to contribute to memory impairments in MDD. Memory did not seem to be impaired in depressed BD, however, SA appeared to contribute to memory performance. Impairments in

PS, attention, and EF were associated with abnormal cortical thickness in the PCC, superior temporal, parahippocampal, right entorhinal and right lateral occipital areas in euthymic BD.

Discussion: PS and SA appear to be impaired in mood disorders, however, more research is needed to investigate the nature of these impairments in BD and MDD in different mood states, controlling for clinical confounds. PS and EF may be primary impairments in euthymic states, whereas SA plays a role in cognitive functioning in depressed states. Our results highlight potentially important relationships between cognitive functions; further research is needed to unravel the precise cognitive profile in each diagnostic group and how this varies between mood states. Core cognitive dysfunction may be associated with abnormal cortical thickness in several brain regions in BD, including some regions implicated in the default mode network. Future research should further explore brain-cognition associations using multivariate analysis and should account for covariance between different brain structural morphological features, such as cortical thickness, cortical volume, and surface area.

ACKNOWLEDGEMENTS

I would firstly like to acknowledge and thank all the volunteers who contributed their time to participate in the projects included in this thesis.

A heartfelt thankyou to my supervisors, Dr Peter Gallagher, Dr Yujiang Wang, and Dr David Cousins, for their support, dedication, and encouragement. I am very grateful to have had supervisors whose doors were always open and who provided space for open discussions and a friendly face, and who have provided mentorship throughout my PhD and beyond.

Thank you to Dr Lucy Asher, Dr Tom Smulders, Dr Yu Guan, and the rest of the Behaviour Informatics team for giving me the opportunity and support needed to pursue this project.

I would also like to thank colleagues who have provided advice, guidance, and invaluable help in workshopping ideas and discussing the project: Dr Peter Taylor and the rest of the CNNP lab, Professor Hamish McAllister-Williams, Tommaso Viola, and our fellow Behaviour Informatics PhD students. Thank you to the Bipolar Lithium Imaging and Spectroscopy Study (BLISS) study team for the help with the BLISS dataset (Carly Flowers and Dr Fiona Smith), and to the four hardworking students for their assistance with the systematic review (Megan Anwyll, Laura Norsworthy, Luke Corbett, and Mia Schultz-Froggatt).

On a personal note, I am grateful to my family and friends, particularly my parents, Karen Little and Jim Little, for their unwavering support throughout my studies and beyond. I'm grateful to the IoN/Henry Wellcome Building crew for bringing an element of fun to all our lunches and drinks; a special thanks to Jen Nacef for being our pub leader. Thank you to Dr Caroline Rae, a true kindred spirit, whose understanding and support has been invaluable. Finally, I'd like to thank Sean Mills for being a constant source of support, encouragement, and fun throughout this whole process.

CONTRIBUTIONS AND FUNDING

This PhD project was generously funded by the Leverhulme Trust PhD studentship, via the Behaviour Informatics PhD programme at Newcastle University.

The datasets described in Chapter 4 were funded by the Stanley Medical Research Institute, the Mental Health Foundation, and the Medical Research Council (MRC).

The Bipolar Lithium Imaging and Spectroscopy Study (BLISS) project described in Chapter 5 was funded by the MRC. Carly Flowers (CF) was responsible for data collection, supervised by Dr David Cousins (DC). Dr Fiona Smith (FS) oversaw the magnetic resonance imaging protocol and acquisition at the Newcastle Magnetic Resonance Imaging Centre.

PUBLICATIONS AND CONFERENCE PRESENTATIONS

The following publications and conference presentations have resulted from the work described in this thesis:

PUBLICATIONS

Little, B. (2022). Future approaches to investigating cognitive functioning in Bipolar Disorder, *Bipolar Disorders*, pp. 1–2.

CONFERENCE PRESENTATIONS

Little B., Wang Y., Cousins D., & Gallagher P. (2022). The Effect of Core Cognitive Dysfunction on Memory in People with Mood Disorders. In the *Annual Mid-Year Barcelona Meeting of the International Neuropsychological Society, Barcelona, Spain*.

Little B., Anwyll M., Norsworthy L., Corbett L., Schultz-Froggatt M., Cousins D., Wang, Y., & Gallagher, P. (2022). Processing speed and sustained attention in mood disorders: A systematic review and meta- analysis. In the *24th Annual Conference of the International Society for Bipolar Disorders, 10–12 June 2022*. p. 33.

CONFERENCE POSTERS

Little B., Gallagher P., Cousins D., Wang Y. (2022). Investigating brain-cognition associations in bipolar disorder using Canonical Correlation Analysis (CCA). In the *Organization for Human Brain Mapping Annual Meeting 2022, Glasgow, UK,* and in the *Federation of European Neurosciences Forum 2022, Paris, France.*

Little, B., & Gallagher, P. (2021). Core cognitive functions and their role in general cognitive dysfunctioning in euthymic bipolar disorder. *In* the *23rd Annual Conference of the International Society for Bipolar Disorders. BIPOLAR DISORDERS (Vol. 23, pp. 65-65)*. NJ, USA: WILEY.

COVID-19 IMPACT STATEMENT

The Covid-19 global pandemic first affected the UK in March 2020, in the first year of this PhD project. The project consequently experienced some setbacks, and we lost time and opportunities to collect additional data. We originally planned to collect MRI and neuropsychological data from adults with MDD (estimated *n*=30), including some neuropsychological tests of interest to this thesis (a Continuous Performance Test for ex-Gaussian analysis of RT data to measure sustained attention and the Hick-Hyman paradigm to measure processing speed). The pandemic delayed this project by over a year and consequently we were unable to collect this data for this thesis. We focussed on the use of extant datasets, but this limited some of our research questions, statistical power, and generalisability of the results. I would like to note the impact this had on the project.

AUTHOR'S DECLARATION

This thesis is submitted to Newcastle University for the degree of Doctor of Philosophy. The research detailed within was performed between the years 2019-2022 and was supervised by Dr Peter Gallagher, Dr Yujiang Wang, and Dr David Cousins. I certify that none of the material offered in this thesis has been previously submitted by me for a degree or any other qualification at this or any other university.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ACC	Anterior cingulate cortex
ADHD	Attention Deficit Hyperactivity Disorder
ANT	Attentional Network Test
BD	Bipolar disorder
BD-d	Bipolar disorder depressed
BD-e	Bipolar disorder euthymic
BDI	Beck Depression Inventory
BIC	Bayesian information criterion
BLISS	Bipolar Lithium Imaging and Spectroscopy Study
CANTAB	Cambridge Neuropsychological Test Automated Battery
CCA	Canonical correlation analysis
CCN	Cognitive control network
CI	Confidence interval
CNSVS	CNS Vital Signs
CoV	Coefficient of variation
CPT	Continuous performance test
CRT	Cognitive remediation therapy
CW	Colour-word (condition of the Stroop test)
df	Degrees of freedom
DMN	Default mode network
DSM	Diagnostic and Statistical Manual of Mental Disorders
DSST	Digit symbol substitution test
EBIC	Extended Bayesian Information Criterion
EEG	Electroencephalogram
EF	Executive function
ES	Effect size
FLAIR	Fluid attenuated inversion recovery
FDR	First-degree relatives
HAM-17	Hamilton Depression Rating Scale 17-item version
HC	Healthy control
ICD	International Classification of Diseases
ISBD	International Society for Bipolar Disorders
iSD	Individual standard deviation

LASSO	Least absolute shrinkage and selection operator
LLD	Late-life depression
MATRICS	Measurement and Treatment Research to Improve Cognition in Schizophrenia
MCCB	MATRICS Consensus Cognitive Battery
MDD	Major depressive disorder
MICE	Multiple imputation by chained equations
MRI	Magnetic resonance imaging
ms	Milliseconds
NART	National Adult Reading Test
PASAT	Paced Auditory Serial Addition Test
PCA	Principal components analysis
PCC	Posterior cingulate cortex
PFC	Prefrontal cortex
PS	Processing speed
PVT	Psychomotor Vigilance Task
RAVLT	Rey Auditory Verbal Learning Test
ROI	Region of interest
RT	Reaction time
SA	Sustained attention
SCID-I	Structured Clinical Interview for DSM-IV-TR Axis I Disorders
SCIP	Screen for Cognitive Impairment in Psychiatry
SCOLP	Speed and Capacity of Language Processing test
SD	Standard deviation
SE	Standard error
SMD	Standardised mean difference
SWM	Spatial Working Memory
TMT	Trail Making Test
ToL	Tower of London
VM	Verbal memory
VS	Visuo-spatial
WAIS	Weschler Adult Intelligence Scale
WM	Working memory
YMRS	Young Mania Rating Scale

CONCEPTS AND DEFINITIONS

Given that there is some heterogeneity in the literature in the use of some terms and definitions, the following terms are defined for the purposes of this thesis.

Mood disorder: While the term 'mood disorder' could be applied to several clinical diagnoses, this thesis focussed on Major Depressive Disorder (MDD) and Bipolar Disorder (BD) as defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM; American Psychiatric Association, 2013) or the International Classification of Diseases (ICD; World Health Organization, 1992) criteria.

Euthymic/remitted: Participants with BD or MDD who did not meet the criteria for a symptomatic (depressive, manic, or mixed) episode at the time of testing. The BD literature generally uses the term 'euthymic' or 'euthymia', whereas the MDD literature generally uses the term 'remitted' or 'remission'; this thesis generally followed this convention, however, both terms were used to denote an absence of symptomatic mood states.

Symptomatic: Participants with BD or MDD who met the criteria for either a manic, depressed, or mixed mood state at the time of testing.

Depression/depressed: This term is often used to describe a clinical diagnosis of MDD, but it can also refer to a discrete event of a depressed mood, which could apply to any person, including healthy controls and other clinical groups. In this thesis, these concepts were distinguished by using 'MDD' to refer to the clinical disorder, and 'depression', 'depressed mood' or 'depressed state' to refer to a discrete event state of depression. MDD is sometimes referred to as unipolar depression to distinguish it from bipolar depression.

Cognitive domain/function: Cognitive abilities. Cognitive domain is often used in the literature as an umbrella term to refer to general cognitive abilities that encompass several distinct functions. For example, the domain of attention may include divided attention, sustained attention, and selective attention. Cognitive function, on the other hand, is often used to refer to individual, distinct cognitive processes. In this thesis, both cognitive functions and domains were conceptualised by scores from neuropsychological tests; these may consist of a single test score or an average of several individual test scores. This thesis generally used the term domain to refer to more general, overarching cognitive abilities. However, we did

not assume a strict distinction between cognitive domains and functions, and we often used the terminology described in previous studies which may have been used interchangeably.

Impairment/deficit/dysfunction: These terms are often used interchangeably in the literature to describe poor cognitive functioning in clinical groups. However, specific terms are sometimes used to denote different concepts, e.g., significantly poorer functioning compared to a control sample, percentile scores, or scores below other thresholds or norms. This thesis focused on comparisons of patients to a control group, so, unless stated otherwise, these terms were used interchangeably to denote significantly poorer functioning in patient groups compared to healthy control groups.

Composite score: A combination of individual test scores that are highly related to one another, e.g., several tests of processing speed averaged to create a composite score for processing speed. They were typically used to represent performance on a cognitive domain.

Grapho-motor tasks: Tests involving pen-and-paper tasks, in which participants are required to complete the task by drawing or writing by hand (as opposed to computerised tasks).

Digit Symbol Substitution Test (DSST): A test of processing speed, traditionally grapho-motor in nature. Several versions of DSST exist, sometimes with different names, such as 'Symbol Digit Modalities Test', but for the purposes of this thesis, we referred to all such types of tests as 'DSST'.

Trail Making Test part A (TMT-A): A test of processing speed, traditionally grapho-motor in nature. Several versions of TMT exist, sometimes with different names, such as 'Colour Trials Test', but for the purposes of this thesis, we grouped all such types of tests as 'TMT-A'.

Continuous Performance Test (CPT): A test of sustained attention. There are different versions of the task that share a similar paradigm. In this thesis, all such tasks were referred to as CPTs, and where possible, the specific outcome measure was specified.

Effect size (ES): ESs measure the magnitude of the difference between two group means. In this thesis, ESs were used to compare a clinical group with a healthy control group on cognitive scores. Unless stated otherwise, ESs reflected Cohen's *d* (the standardised mean difference), and were interpreted using on the following criteria: ESs ranging 0.2-0.5 were considered small; 0.5-0.8 were moderate ESs; and >0.80 were considered large ESs (J. Cohen, 1988).

Chapter 1 INTRODUCTION TO COGNITIVE FUNCTIONING IN MOOD DISORDERS

1.1. MOOD DISORDERS

Mood disorders such as Major Depressive Disorder (MDD) and Bipolar Disorder (BD) are characterised by emotional, cognitive, and physical symptoms that impair an individual's ability to function in daily life (Solé et al., 2018). Mood disorders are common, with approximately 9% of the population suffering from them, and MDD and BD are the most commonly diagnosed mood disorders (Fineberg et al., 2013; Malhi et al., 2015). MDD and BD bear a significant burden on individual functioning, healthcare systems and society, and are a leading cause of disability world-wide (Bromet et al., 2011; Fineberg et al., 2013; Hirschfeld & Vornik, 2005; A. H. Young et al., 2011). People with mood disorders tend to suffer chronic illness or have recurrent patterns of the illness (Grande et al., 2016). Researching the nature of mood disorders is therefore vital to improve diagnosis and treatment and lessen the burden on society and the individual.

MDD and BD are typically diagnosed according to criteria set by the International Classification of Diseases 10th revision (ICD-10; World Health Organization, 1992) and Diagnostic and Statistical Manual of Mental Disorders 5th edition (DSM-5; American Psychiatric Association, 2013). These diagnostic manuals characterise MDD as symptoms such as persistent depressed mood, feelings of hopelessness and worthlessness, anhedonia (i.e., loss of interest and pleasure in activities), and loss of energy (American Psychiatric Association, 2013). BD is characterised by periodic changes in mood that vary between periods of depression, mania (i.e., elevated mood, irritability, restlessness and hyperactivity, and racing thoughts that disrupt daily functioning) or hypomania (symptoms of mania that do not disrupt daily functioning), and euthymia (i.e., no mood disturbances) (American Psychiatric Association, 2013). Patients can also experience mixed states, where both symptoms of mania and depression are present (Grande et al., 2016). BD can be divided into type I (BD-I) and type II (BD-II): BD-I is associated with periods of mania, while BD-II is associated with hypomania (American Psychiatric Association, 2013; Angst et al., 2013). BD-I is more common than BD-II (0.6% and 0.4% lifetime prevalence, respectively; Merikangas et al., 2011). The main

difference between MDD and BD is therefore the presence of (hypo)mania in BD, and that MDD tend to experience longer durations of episodes compared to BD (Malhi et al., 2015).

However, there is considerable overlap in these disorders, which can make diagnosis difficult: BD patients are sometimes first diagnosed with MDD before a subsequent diagnosis of BD-II or BD-I after presenting with hypomanic and/or manic episodes (Malhi et al., 2015; Oliveira et al., 2021). The rate of conversion from MDD to BD in 3 years has been estimated at 12.4% (Oliveira et al., 2021). The disorders overlap in symptoms such as depression, cognitive dysfunction, as well as somatic symptoms such as psychomotor disturbance, sleep disturbances, and changes in appetite (Barlow et al., 2016; Malhi et al., 2015; Miskowiak et al., 2012). BD and MDD are both heterogeneous disorders, with patients varying in symptoms and trajectories, which also makes it difficult to distinguish the disorders (Wardenaar & de Jonge, 2013). Given these similarities, it is important to consider both unipolar and bipolar disorders when researching mood disorders, as studying their similarities and differences may be useful for improving diagnosis and treatment.

1.2. COGNITIVE DYSFUNCTION IN MOOD DISORDERS

MDD and BD are both associated with widespread cognitive impairment, with patients showing poorer performance on a broad range of neuropsychological tests compared to healthy controls (HCs) (Ahern & Semkovska, 2017; Cullen et al., 2016; Robinson et al., 2006; Semkovska et al., 2019). Cognitive dysfunction is apparent in symptomatic states (i.e., depression and (hypo)mania) as well as in euthymia, thus it may be considered a core feature of mood disorders (Bo et al., 2017; Bora et al., 2013; Bourne et al., 2013; Malhi et al., 2007; Semkovska et al., 2019). Cognitive impairment appears to be related to poorer general, social, and occupational functioning and quality of life in BD (Baune & Malhi, 2015; Cotrena et al., 2016) and MDD (Ebert et al., 2017; Haro et al., 2019; Withall et al., 2009; Zazula et al., 2021). Longitudinal studies suggest that cognitive functioning has a causal outcome on general functioning in BD (Ehrminger et al., 2019). Thus, investigating the nature of this cognitive impairment may contribute to improving general functioning and daily lives of patients.

Cognitive impairment appears to span most cognitive domains, including processing speed, attention, working memory, executive function, verbal and non-verbal memory, visuospatial function, and language, with the magnitude of the difference compared to healthy controls

ranging from moderate to large effect sizes (Mann-Wrobel et al., 2011; Porter et al., 2015; Zazula et al., 2021). While impairments are found across most cognitive functions in mood disorders, the extent to which this reflects a broad cognitive dysfunction, or whether there may be specific impairments in particular cognitive domains, is not clear (Samamé et al., 2017). When comparing patients to controls, processing speed, executive function and attention often show the largest effect sizes, however, findings are not consistent and significant heterogeneity is found across studies (Ahern & Semkovska, 2017; Bo et al., 2017; R. Lee et al., 2014). More research is needed to establish whether mood disorders are associated with a general or specific cognitive impairment.

Whether or not BD and MDD have distinct cognitive profiles is also unclear. Some studies suggest that BD show a greater magnitude of deficits than MDD, but that BD and MDD do not differ in their pattern of cognitive profile (Cotrena et al., 2016; R. Lee et al., 2015; MacQueen & Memedovich, 2017; Porter et al., 2015; Zazula et al., 2021). In line with this, more severe cognitive impairment in BD appears to be associated with poorer general functioning and quality of life than MDD (Zazula et al., 2021). BD and MDD do not appear to differ in subjective cognitive dysfunction after hospital discharge (Miskowiak et al., 2012). Cluster analysis of cognitive performance has also failed to differentiate BD from MDD, suggesting cross-disorder overlap (Cotrena et al., 2017). However, there are contradictory findings: some studies suggested that unmedicated MDD patients had a broader range of cognitive impairments than BD patients (Tavares et al., 2007). A recent meta-analysis found no significant differences between BD and MDD on cognitive performance other than on tests of verbal memory, however this may be due to methodological differences (Samamé et al., 2017). Disentangling whether there is a difference between the cognitive profile of BD and MDD would help to inform the best approach for future research and could improve clinical practice.

While cognitive impairment is generally found in all mood states, including in euthymia and remission, studies suggest that the impairment may be more widespread and more severe in symptomatic states (Ahern & Semkovska, 2017; Kurtz & Gerraty, 2009; Tsitsipa & Fountoulakis, 2015). For example, greater severity of depressed mood has been associated with greater deficit in memory (Porter et al., 2015). Depressed BD appears to be associated with a higher prevalence of cognitive impairment than euthymic BD (Douglas et al., 2018). Examining the cognitive profile of mania is challenging as behavioural symptoms of mania make recruitment and data collection difficult, however studies suggest that number of manic

episodes is correlated with greater cognitive dysfunction (X. Lin et al., 2019) and patients with current mania show impairments in more functions than depression (Sweeney et al., 2000). Whether symptomatic states have distinct cognitive profiles compared to euthymia, rather just increased magnitude of impairment, is still unclear. Understanding potential differences in cognitive profile between mood states would help us to understand the nature of cognitive impairment in MDD and BD, highlighting whether specific impairments are a trait of the disorder or dependent on mood state, and could perhaps reveal cross-disorder overlaps.

1.3. COGNITIVE HIERARCHY

A complication in this field of research is the way in which studies conceptualise cognitive functions. Research tends to separate cognitive processes into discrete functions or domains, for example the DSM-5 describes six cognitive domains, each encompassing several functions within it, as shown in Figure 1 (Sachdev et al., 2014). Note that this classification was designed for neurocognitive disorders (e.g., delirium and dementia) and no such classification system currently exists for psychiatric disorders. While this approach allows researchers to test specific functions in participants, studies currently differ in the cognitive functions and/or domains that they assess and how they define them, making research difficult to compare. Further, this approach may lead researchers to assume that cognitive functions operate independently as discrete functions, which may be misleading.

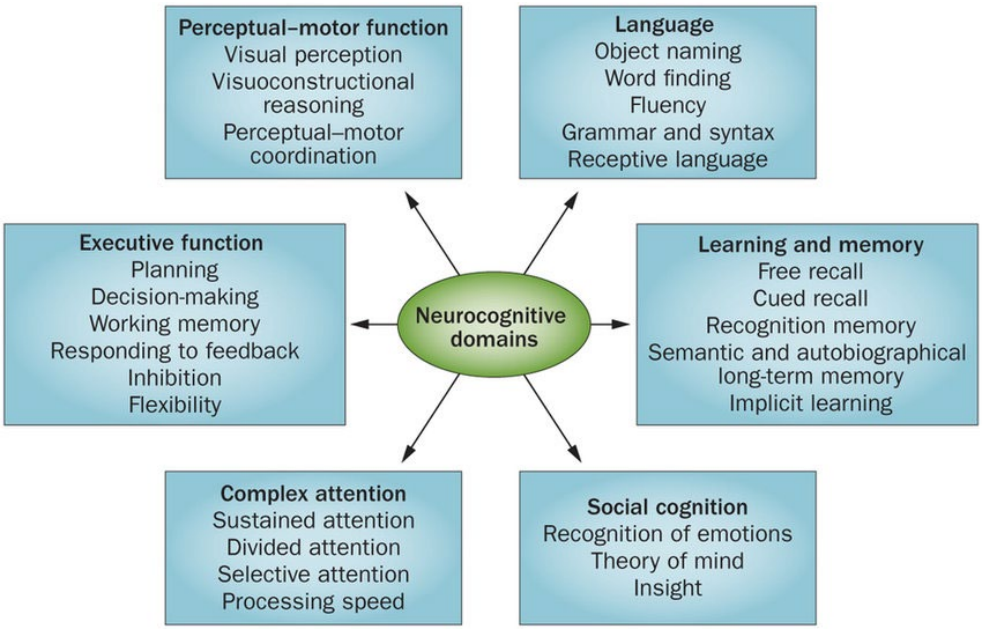

Figure 1: Six cognitive domains and subdomains for classifying neurocognitive disorders as defined by the Diagnostic and Statistical Manual of Mental Disorders Fifth edition (image taken from Sachdev et al., 2014).

In healthy controls, many theoretical models of cognitive processes account for relationships between cognitive functions, allowing functions to overlap or interact. For example, processing speed has been proposed as a fundamental process in healthy cognitive models (Kail & Salthouse, 1994). Processing speed and attention are considered in models of executive function, where functions are considered discrete but interrelated in order to enable executive control (Anderson, 2002). In this sense, we might expect abilities in some cognitive functions have an influence on wider cognitive functioning. For example, performance on a test of memory requires a participant to focus on the task, demanding attention, and faster processing speed would facilitate better performance on most tests of cognitive function. This perspective suggests the existence of a cognitive hierarchy, where some 'core' cognitive functions play a bottom-up or top-down role in higher-order cognition.

Three candidates for being considered 'core' cognitive functions in their potential to influence wider neuropsychological performance are processing speed (PS; also referred to as psychomotor speed), sustained attention (SA), and executive function (EF). For example, with slower PS and reduced ability to maintain attention, an individual may not perform well on any neuropsychological test. Indeed, in healthy people, PS has been shown to affect a broad range of other cognitive processes, such as working memory, visuospatial processing, language abilities, and EF (Finkel et al., 2005; Gilsoul et al., 2019; Jakobsen et al., 2011). PS is thought to be related to general intelligence and is consequently included in several intelligence test batteries (Kail & Salthouse, 1994; Lange & Lippa, 2016; Schubert et al., 2017; Weshsler, 1997). Age-related decline in PS has also been theorised to account for healthy cognitive ageing (Salthouse, 1996). However, the extent to which it is related to intelligence and other cognitive domains is debated (Conway et al., 2002; Sheppard & Vernon, 2008). EF has similarly been shown to account for a large amount of variance in memory performance in healthy adults (Duff et al., 2005). Establishing whether specific functions contribute to general functioning is important in our understanding of human cognitive functioning as well as for researching where these processes deviate from the norm in the context of pathology.

1.3.1. Cognitive hierarchy in mood disorders

Given that cognitive functions are likely interrelated in healthy models, it follows that if one or more functions become impaired via pathological mechanisms, e.g., in psychiatric disorders, this could have a knock-on effect for wider functioning. Some theories of the

mechanisms behind cognitive dysfunction in mood disorders suggest that specific functions may be responsible for more general impairments: the cognitive speed hypothesis states that cognitive impairment in MDD is characterised by slower processing, and has been supported by studies that found an impairment on tests of PS, but not on tasks requiring effortful EF (den Hartog et al., 2003; Hu et al., 2022; Nebes et al., 2000). Alternatively, the cognitive effort hypothesis states that MDD is characterised by problems with allocating effort to cognitive tasks, where effortful processing can be defined as activities such as rehearsal, imagery, organisation or systematic searching (Austin et al., 2001; den Hartog et al., 2003). An example of these two processes can be found in the Stroop task, where simple trials (word reading and colour naming) are considered automatic processes, whereas the colour-word interference task requires more effort to inhibit the automatic word reading response to read the colour of the font. The cognitive effort hypothesis suggests that performance on tasks requiring more cognitive effort will be disproportionately impaired compared to automatic processes in MDD. However, the potential roles of automatic processing and effortful processing in depression-related cognitive impairment are not necessarily mutually exclusive: it is possible that slower processing speed contributes to wider cognitive function in a bottom-up way, whereas problems with allocating effort to tasks, involving elements of EF, may simultaneously have a top-down effect on performance.

1.3.2. Cognitive hierarchy in bipolar disorder

Researching the nature of human cognition is a challenge, as cognitive functions are latent variables that are measured indirectly with neuropsychological tests. Dimension reduction techniques have been used to investigate the latent cognitive profile of people with mood disorders: Gallagher et al. (2014) applied Principal Components Analysis (PCA) to neuropsychological data from depressed BD patients and HCs to investigate the underlying cognitive factor structure. In HCs, four cognitive components were found: visuo-spatial (VS) short-term/immediate processing; VS self-ordered/strategic processing; verbal learning and memory (VM); and verbal EF and working memory (WM). The cognitive components in the BD group were qualitatively similar to controls, but there were only 3 components (executive control and VS memory, VM, and a component containing VS memory and verbal WM tasks), suggesting overlap between cognitive functions in BD. Cognitive processes perhaps function less independently in BD depression than in a healthy model, which may suggest the presence of cognitive scaffolding in this group. A follow-up hierarchical regression analysis of this data

showed that verbal learning explained a significant amount of variance in VS memory and accounted for the between-group difference, suggesting that patients engaged verbal learning processes to support diminished VS resources (Gallagher, Gray, et al., 2015). This outlines the importance of accounting for interrelationships between cognitive functions.

Many studies have concluded that there exists a generalised rather than specific cognitive impairment in mood disorders because they did not find evidence that any cognitive function was particularly impaired (Hill et al., 2009; Majer et al., 2004; Mann-Wrobel et al., 2011). However, most studies do not directly test the nature of relationships between cognitive functions, so cannot conclude whether specific cognitive functions may have caused the seemingly general impairment. Mediation and hierarchical regression models have been utilised to investigate such relationships: a twin study showed that PS mediated memory impairment in BD, accounting for a large part of the variance in WM, VM, and VS memory in both BD patients and their first-degree relatives (FDRs) (Kieseppä et al., 2005). This was shown using both a graphomotor test of PS (the Digit Symbol Substitution Test; DSST), as well as computerised reaction time (RT). Antila et al. (2011) similarly found that DSST performance accounted for a large part of the variance in memory in BD and FDRs; the role of PS was least extensive in controls, more so in relatives, and most extensive in patients, where it accounted for most cognitive deficits. Motor speed has also predicted BD-related deficits in verbal fluency, reasoning, EF, attention, and general cognitive function (Salazar-Fraile et al., 2009). EF likewise appears to explain memory impairments in BD: while patients showed impairments across all cognitive domains, hierarchical regression analysis has shown that EF and PS together can account for the between-group variance in memory, even after controlling for mood severity (Thompson et al., 2009). However, this study did not consider SA in their analysis, and instead grouped tests of attention (e.g., Vigil Continuous Performance Test errors) with the EF domain. Nevertheless, together, these studies suggest that poor PS and EF may partly explain memory impairments in BD.

1.3.3. Cognitive hierarchy in major depressive disorder

Similar results have been found for people with unipolar depression: PS and EF appear to affect memory and language in late-life depression (LLD) (Butters et al., 2004; Nebes et al., 2000; Sexton et al., 2012; Sheline et al., 2006). However, most research to date tested this concept in older adults, where the depression-related effects may be confounded by the

presence of normal cognitive ageing (Butters et al., 2004; Formánek et al., 2020). One recent study of adults with MDD reported that PS mediated the relationship between depression status and VM and VS memory, even after controlling for age, sex and premorbid IQ (Zaremba et al., 2019). Liu et al. (2019) measured PS, EF, attention, and memory in MDD and controls, and used hierarchical regression models to assess relationships between the cognitive functions. After controlling for demographic variables and the other cognitive domains, the effect of group on PS and EF remained significant, suggesting these were primary impairments. Some contradictory research exists, however, with some studies failing to find evidence that PS can explain general cognitive impairment in MDD (McDermott & Ebmeier, 2009). However, this conclusion was made because timed and untimed tests were similarly related to depression severity and no formal comparisons were made of the strength of correlations or interrelationships between cognitive functions. The literature generally suggests that there may primary cognitive impairments in MDD that lead to secondary dysfunction, however, the full picture is not yet clear.

While EF may affect wider dysfunction, some research suggests that EF may be secondary to attentional impairment in MDD (Nilsson et al., 2016), thus, poor attention may also play a role in depression-related cognitive deficits. However, this study did not consider SA specifically, conceptualising attention more generally. There was no cognitive domain representing PS and instead some tasks that may be considered measures of PS were included in the attention domain (e.g., the Trail-Making Test part A and Stroop automatic trials), complicating interpretation of their results. Poor performance on tests of EF in MDD may also be caused, at least in part, by slower processing: PS and motor speed appear to explain impairments in the Stroop task in MDD, suggesting that impairment on this task, thought to measure EF and attention, may be due to psychomotor slowing (Kertzman et al., 2010). Slower PS is also thought to underlie verbal fluency impairments in MDD (Henry & Crawford, 2005). A meta-analysis found that PS could not completely account for the impairment in EF in MDD patients (Snyder, 2013); core cognitive functions may therefore be related, but their individual influence on general cognitive impairment in MDD is yet to be established. This research hints at a hierarchy of cognitive dysfunction in BD and MDD, where impaired core functioning has a detrimental effect on wider cognitive performance. Investigating the nature this cognitive hierarchy is vital for unravelling the mechanisms underpinning cognitive dysfunction in mood disorders and could help inform cognitive remediation interventions.

1.4. CORE COGNITIVE DYSFUNCTION IN MOOD DISORDERS

Since PS, SA and EF may have important roles in cognitive dysfunction in BD, it is important to firstly establish the magnitude of these impairments in BD and MDD. The finding that EF is impaired in mood disorders is relatively well-established (Cotrena et al., 2020; Goswami et al., 2006; Normala et al., 2010; Snyder, 2013). EF can be split into three dissociable sub-types: inhibition, shifting and updating WM (Miyake, Friedman, et al., 2000). Studies suggest that BD and MDD do not show a differential deficit on any of these subcomponents and instead display impairments in all aspects of EF (Henry & Crawford, 2005; Thompson et al., 2009). A multivariate analysis found that tests of tests of EF, including set shifting, inhibition, and fluency, had strongest discriminating ability in BD vs controls (Sparding et al., 2015). EF has also been linked to long-term functional outcome in BD (Bonnín et al., 2010). Less research exists on the nature of impairments in PS and SA in BD and MDD.

Slower PS in BD and MDD compared to controls has been reported in previous studies and meta-analyses (Kriesche et al., 2022; J. Liu et al., 2019; Luperdi et al., 2021; Semkovska et al., 2019). Some research suggests that PS may be particularly impaired in mood disorders, with tests of PS showing the largest ESs when comparing patients to controls on a range of cognitive tests (Gallagher et al., 2014; Thompson et al., 2005). Slow PS has been found in euthymic BD patients as well as unaffected FDRs of patients, even after controlling for current symptom severity, premorbid IQ, and education, suggesting it may be an endophenotype for BD (Daban et al., 2012; Dobri et al., 2022; Fears et al., 2014; Luperdi et al., 2021). A recent study found that PS and memory, but not EF or attention, were impaired in first-episode MDD, after controlling for confounding demographic variables (Hu et al., 2022). However, research into the breadth and degree of psychomotor slowing in mood disorders is not consistent, with some studies failing to find an impairment in PS in mood disorder groups (Porter et al., 2003; Rock et al., 2014).

Attentional impairments in have been found in patients with BD and MDD (Camelo et al., 2013; X. Wang et al., 2020). A recent study used a go/no-go task to measure different attentional processes in MDD, including focussed attention (errors of omission), response inhibition (commission errors), alertness (RT), and sustained attention (variability in RT) (Schmidt et al., 2021). While MDD showed impairments on all four functions compared to controls, SA was the strongest predictor of MDD, suggesting that SA is the most affected

attentional process in MDD (Schmidt et al., 2021). Similarly, a recent meta-analysis showed that the majority of studies found that MDD performed worse than HCs on tests of attention, particularly for tests of alertness and SA (Kriesche et al., 2022). People with BD also appear to show poorer SA than HCs, even in euthymia (Ancín et al., 2010; Gallagher, Nilsson, et al., 2015). The impairment in SA does not appear to be due to poor WM, thus may reflect a core feature of BD (Harmer et al., 2002). Impairments in SA are found in euthymia (Ancín et al., 2010), and FDRs of people with BD also show impaired EF and SA, suggesting it may be an endophenotype of the disorder (Miskowiak, Kjærstad, et al., 2017). However, contradictory results exist, with not all studies detecting impaired attention or SA in BD and MDD (Gallagher, Nilsson, et al., 2015; Goswami et al., 2006; Maalouf et al., 2010).

Mood disorders may therefore be associated with impairments in PS and SA compared to healthy controls. These impairments appear to relate to general functioning, for example, psychomotor slowing was related to worsening of functional disability in a longitudinal study of MDD and BD (R. Lee et al., 2015) and PS appears to be related to social and global functioning in BD (Burdick et al., 2010). PS and attention also appear to predict later social impairment 18 years later in BD and MDD, even after controlling for baseline social functioning and depressive symptoms (Sarapas et al., 2013). Improved SA has likewise been associated with general functioning including employment in a longitudinal study of MDD and BD (R. Lee et al., 2015). Impairments in these functions may also be unique to mood disorders, as patterns of SA and PS may differentiate mood disorder groups from other psychiatric group such as schizophrenia (Bora et al., 2009). PS and SA may therefore be of particular importance to study and may have effects on general functioning of patients, however more research is needed to uncover the precise nature of these impairments.

1.5. Neural correlates of cognitive dysfunction in mood disorders

To gain the full picture of core cognitive dysfunction in mood disorders, potential associations of cognitive dysfunction with brain abnormalities should be explored. While many studies have investigated brain abnormalities and how they relate to cognitive dysfunction in mood disorders (Hanford et al., 2016; Jamieson et al., 2019), the complex neurobiological underpinnings of BD and MDD are not well understood. There exists a wealth of literature on this topic; we did not aim to provide an exhaustive review here, but the following sections

highlight that there are potential relationships between brain structure and cognitive functions of interest to this thesis.

1.5.1. Brain abnormalities in mood disorders

BD appears to be associated with structural abnormalities in the brain: studies have found decreased cortical thickness in BD patients compared to HCs across widespread areas of the brain, even in euthymia, which may reflect abnormal tissue microstructure (Foland-Ross et al., 2011; Macoveanu et al., 2021; Necus et al., 2021). BD patients appear to have a lower volume of subcortical regions than controls, particularly in the hippocampus and thalamus (Hibar et al., 2016). Gyrification also seems to be affected in BD, with patients showing reduced gyrification in the pre-frontal cortex (PFC) compared to HCs (A. McIntosh et al., 2009). Structural connectivity is likewise altered in BD, with patients showing widespread white matter abnormalities, reduced white matter integrity, and abnormal myelination compared to HCs (Favre et al., 2019; Kempton et al., 2008; Lloyd et al., 2009; Masuda et al., 2020; Necus et al., 2019, 2021; Wise et al., 2016). Cortical and subcortical volume, cortical thickness, surface area, and white matter abnormalities are similarly found in MDD (Lemke et al., 2022; Schmaal et al., 2016, 2017; van Velzen et al., 2020). Studies have indicated that patterns of cortical thickness may differ between BD and MDD (Lan et al., 2014), suggesting that different regions are implicated in each disorder. BD and MDD also appear to show abnormalities in functional connectivity and resting-state brain functioning compared to controls (Chen et al., 2018; Y. Wang et al., 2015). While it appears that structural and functional abnormalities are common in mood disorders, there is not yet a consensus of which regions are particularly affected as results are heterogeneous (Hanford et al., 2016; Z. Zhu et al., 2022).

1.5.2. Brain-cognition associations

A link between cognitive functioning and brain structure in healthy controls is well established (Deary et al., 2010; Dickerson et al., 2008), and cognitive performance has been related to specific structural features of the brain, including cortical thickness, cortical volume, surface area, and gyrification (Gautam et al., 2015; Hartberg et al., 2011; Walhovd et al., 2016; Zimmerman et al., 2006). In terms of core cognitive functioning, brain morphological metrics such as cortical thickness have been linked to PS, attention, and EF in healthy adults (Pol et al., 2006; Weise et al., 2019; Westlye et al., 2011). Given the apparent brain-cognition

associations in healthy people, we might then expect impaired cognitive functioning to be related to abnormal brain structure in psychiatric groups.

Brain structural abnormalities have indeed been linked to poorer cognitive functioning in MDD (Geraets et al., 2021; Jamieson et al., 2019). Abnormal cortical thickness and surface area have also been associated with cognitive impairment in BD (Hartberg et al., 2011; Hatton et al., 2013; Kang et al., 2022). Reduced white matter integrity appears to be related to cognitive dysfunction in BD and MDD (Kieseppä et al., 2014; Rizk et al., 2017). However, findings are not consistent, with some studies failing to find evidence for a relationship between brain structure and cognitive functioning (Alonso-Lana et al., 2016; Gutiérrez-Galve et al., 2012). Studies of the association between cortical morphology and cognitive function in BD and MDD are scarce (Jamieson et al., 2019; Karantonis, Carruthers, et al., 2021) and further research is needed to assess brain-cognition associations in each group.

Core cognitive functions appear to be related to abnormal brain structure in mood disorders: for example, PS appears to be related to surface area and cortical thickness in BD (Fears et al., 2015; Hartberg et al., 2011; Oertel-Knöchel et al., 2015). SA has likewise been related to cortical thickness and cortical and hippocampal volumes in BD (Hatton et al., 2013; Sax et al., 1999). EF has also been associated with cortical thickness, volume, and white matter integrity in BD (Abé et al., 2018; Hartberg et al., 2011; Poletti et al., 2015). However, some studies failed to find an association between cortical thickness and core cognitive functions in patients (Gutiérrez-Galve et al., 2012; J. U. Kim et al., 2022; Knöchel et al., 2016). Overall, research hints at associations between core cognitive dysfunction and abnormal brain structure in mood disorders, but few studies have measured brain-cognition associations in patients, especially those that include measures of core cognitive functioning.

1.5.3. Brain function and cognitive impairment

People with mood disorders also tend to show abnormalities in brain function that relate to cognitive impairment: functional magnetic resonance imaging (fMRI) studies show that impairments in several cognitive domains are associated with abnormal activity in the PFC and anterior cingulate cortex (ACC) (Mesbah et al., 2023; Miskowiak & Petersen, 2019; Piani et al., 2022). For example, poor sustained attention in BD has been linked to hypo-activation of regions associated with cognitive control, including the dorso-lateral PFC, ventro-lateral PFC, and parietal cortex (Fleck et al., 2012; Smucny, Lesh, Newton, et al., 2018). Abnormal

functional connectivity of the Default Mode Network (DMN), a network of brain regions that are active during resting states, has been related to cognitive dysfunction in BD (Miskowiak & Petersen, 2019; Nguyen et al., 2017). Compared to HCs, people with BD show hypoconnectivity of the DMN (Meda et al., 2014), as well as less variable temporal connectivity strength between regions in the DMN (medial PFC and posterior-cingulate cortex [PCC]), where a higher degree of this abnormality has been linked to slower PS and poorer EF (Nguyen et al., 2017; Rashid et al., 2014). Similarly, BD show abnormal functional connectivity in the DMN during resting-states, particularly in links between the medial PFC and PCC, and this has been associated with impairments in EF (Massalha et al., 2023). BD patients therefore appear to show a failure to deactivate the DMN and insufficient recruitment of task-relative cognitive control regions, which are associated with worse cognitive performance (Miskowiak & Petersen, 2019; Zarp Petersen et al., 2022).

While there is some heterogeneity in findings of associations between brain function and cognition in mood disorders (Piani et al., 2022), results point towards a role of PFC activity and DMN-related activity (Macoveanu, Petersen, et al., 2023; Miskowiak & Petersen, 2019). Recent research suggests that cognitive interventions that modulate activity of the DMN and cognitive control network (CCN) are associated with cognitive improvements (Miskowiak, Yalin, et al., 2022). Functional imaging may therefore detect brain abnormalities that are more closely linked to cognitive impairment in real time than structural metrics, where there are less consistent findings in identifying key regions. While fMRI offers an insight into how regions and networks are dynamically recruited during cognitive tasks, and how this may differ in patients with cognitive impairment, structural abnormalities may reflect more steady physical features or changes associated with cognitive impairment, that may be a biological trait of the disorder or may progress over time. Together, functional and structural imaging studies can provide complementary information on the nature of brain-cognition associations in mood disorders.

1.5.4. Interim summary

The literature suggests that mood disorders are associated with cognitive impairments, including core cognitive impairments that may influence wider functioning, however the specific cognitive profiles of BD and MDD are not yet clear. How these core cognitive functions relate to structural brain abnormalities in mood disorders is also unclear, with mixed results

in the literature. Heterogeneous findings may have resulted from various methodological approaches used across studies. Chapter 2 describes some of these methodological issues and highlights approaches that may improve research in this field.

Chapter 2 METHODOLOGICAL APPROACHES TO INVESTIGATING COGNITIVE FUNCTIONING IN MOOD DISORDERS

The content of this chapter has been summarised and published in a commentary article (Little, 2023).

2.1. NEUROPSYCHOLOGICAL METHODOLOGY

While there appear to be impairments in PS, SA, and EF in mood disorders that may disrupt wider functioning, findings are not consistent and significant heterogeneity is found across studies (Ahern & Semkovska, 2017; Bo et al., 2017; R. Lee et al., 2014). Mixed results may be driven, at least in part, by heterogeneous methodology in the literature, including differences in neuropsychological assessment across studies (Bourne et al., 2013; Cardenas et al., 2016). Attempts have been made to standardise neuropsychological assessment by recommending the use of validated test batteries in research and clinical practice, such as the Screen for Cognitive Impairment in Psychiatry (SCIP), the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognitive Battery (MCCB), or other standardised batteries or tests (Douglas et al., 2018; Miskowiak et al., 2018; Miskowiak, Burdick, et al., 2017; Purdon, 2005; Yatham et al., 2010). However, there is not yet a consensus across researchers (MacQueen & Memedovich, 2017). The need for test batteries to be brief in clinical practice means there are specific cognitive functions such as PS and SA that are sometimes missed by test batteries (McIntyre et al., 2017; Nasreddine et al., 2005). Neuropsychological batteries therefore benefit from the use of standardised, replicable instruments, but may lack breadth of cognitive tests.

2.1.1. Tests of processing speed and sustained attention

In terms of specific neuropsychological tests used to measure PS and SA, there are some commonly used, standardised instruments. For example, the Trail Making Test (TMT) is often used to measure PS, attention, and EF (Misdraji & Gass, 2010; Reitan, 1958). In part A of TMT, circled numbers are randomly distributed on a sheet of paper and participants must connect numbers in ascending sequential order by drawing lines between the circles (Meyers, 2017; Reitan, 1958). The outcome variable used is typically the time taken to complete the test, but number of errors are sometimes recorded. Digit-symbol substitution tests (DSST) are commonly used to test PS in the mood disorder literature (Semkovska et al., 2019). DSSTs

require subjects to transcribe a series of unique geometric symbols using a key that displays corresponding Arabic numbers (1-9). The outcome variable is typically the number of correct responses in a set time (e.g., 90 seconds) but can sometimes be the time to complete the task. There are several different versions of this task, which are often included in various neuropsychological test batteries, for example the DSST is a subtest of the Wechsler Adult Intelligence Scale (Weshsler, 1997); the versions may vary in materials, e.g., computerised or graphomotor, but share the same paradigm. PS can also be measured with timed verbal and language tasks (Baddeley et al., 1992; Henry & Crawford, 2005). For example, in verbal fluency tasks, participants name as many words as possible starting with a specific letter (phonemic fluency) or from a specific category (category fluency). The outcome measure is usually the number of appropriate words stated in a set time (e.g., 60 seconds), but sometimes the number of errors are reported. This test is often used to measure EF, but since it is a timed task, it is sometimes used to measure PS and has been suggested to reflect a deficit in PS, rather than EF, in mood disorders (Henry & Crawford, 2005).

PS can also be measured with reaction time (RT) paradigms (A. Jensen & Munro, 1979). These are usually conceptualised as either simple-RT tasks, where participants must react as soon as they see a target stimulus presented alone, or choice-RT tasks, where participants must react to a target among distractor stimuli. RT is usually defined as the time taken for a participant to respond to a target stimulus. The Stroop task is sometimes used to measure PS: in this task, participants must read aloud the colour of the ink of a word presented on screen and ignore the word itself (Stroop, 1935). When the colour of the ink is incongruent with the word, this is thought to measure EF (inhibition). However, in simple/automatic trials participants either name a block of colour presented or read words aloud; scores from these trials can be used to measure PS since they are similar to simple-RT tasks. The outcome measure is either the time take to complete a given number of trials or the number of trials correctly completed in a set time. A variety of tests of PS therefore exists, with a range of tests being used to represent PS in the mood disorder literature (Semkovska et al., 2019).

Continuous Performance Tests (CPT) are typically used to measure SA and involve the presentation of a sequence of stimuli (usually letters or numbers) over an extended period. The participant must respond to a specific target (e.g., pressing a button when they see a particular letter) as quickly as possible, ignoring distractors (R. Cohen, 2019; Conners et al., 2003). The task therefore involves a speed accuracy-trade off (Fitts, 1954). There are several

versions of the CPT paradigm, with varying durations, types of stimuli, and target frequencies (Moss et al., 2016), such as the Vigil CPT (Cegalis & Bowlin, 1991). Several outcome variables can be taken from this test, including accuracy measures (correct/hits and errors of omission and commission), RT, RT variability, and other models of RT such as ex-Gaussian analysis. Studies vary in the chosen outcome measure from CPTs to represent SA, with some using RT-based measures and others using accuracy measures (Bo et al., 2017; Bora et al., 2006). Other measures of vigilance, such as the Psychomotor Vigilance Task (PVT) have a simple-RT design, requiring subjects to respond to a single stimulus which appears at random intervals over an extended period of time (Arsintescu et al., 2019).

2.1.2. Test validity and the task impurity problem

An issue restricting such research involves the validity of neuropsychological assessment. A difficulty here lies in our ability to disentangle distinct cognitive functions from one test score, often termed the *task impurity problem* (Miyake, Emerson, et al., 2000). For example, PS can be defined as the speed in which an individual processes and reacts to stimuli. It is therefore usually conceptualised using scores from time-sensitive tasks, as mentioned above. However, the entire process from perceiving the stimulus to responding comprises several components, which rely on more than just PS: first, the participant must perceive the stimulus and process its properties, then they must understand the task and decide how to react, before initiating the behavioural response. The process therefore involves a combination of sensory, perceptual, cognitive, decision-making, and motor components and the exact nature of this process is not well understood. Tests of PS have been shown to require elements of EF, where the demand on EF varies between tasks (Cepeda et al., 2013), demonstrating how the task impurity problem can confound results.

Attention can be described as the ability to direct resources to particular stimuli that may be associated with a goal (R. Cohen, 2014). Attention has several sub-categories or manifestations, including sustained, divided, and selective attention, which all facilitate cognitive performance (Gunstad et al., 2006). SA, or vigilance, characterises the ability to concentrate on a specific stimulus and ignore irrelevant stimuli over a period of time, and is therefore related to general alertness and mind-wandering (R. Cohen, 2014; Esterman & Rothlein, 2019). Given that PS is measured with simple cognitive tasks that will require SA, and that SA tasks would be facilitated by faster processing, disentangling these cognitive

components proves challenging and there is little research that attempts to investigate each function in closer detail in mood disorder populations. Differentiating the various perceptual, cognitive, and motor components would help in furthering our understanding of cognitive hierarchy and how it may be impaired in clinical populations.

It may be impossible to design a 'pure' measure of a specific cognitive function, but statistical efforts can be made to mitigate the task impurity problem and to unravel components from messy multicomponent data. For example, in the case of the DSST, subtests of the task (e.g., Symbol Copy) can help fractionate cognitive components. In the Symbol Copy variant, subjects directly copy geometric symbols, rather than coding them from a key, which aims to tests motor speed by removing the elements of visual scanning and memory. Subtracting this score from the original DSST score will remove the motor time, theoretically leaving a measure that reflects only cognitive speed (Joy, Fein, et al., 2003; Joy, Kaplan, et al., 2003). Similarly, several tests of EF contain similar tasks that require more automatic processing, which can be subtracted from the more effortful task score: in the case of the Stroop task, the automatic word-reading or colour-naming scores can be subtracted from the colour-word score; for digit span, forward span can be subtracted from backwards span; and for the TMT, part A can be subtracted from part B. While such strategies are useful in picking apart separate cognitive functions, they are not available for all neuropsychological tests and are arguably under-utilised.

2.1.3. Modelling reaction time

Similarly, statistical models can be used to conceptualise SA more precisely. Studies often use the average RT from CPTs to measure SA, however, RT data has several issues. Firstly, RTs encompass PS as well as attention, so it is not a pure measure of SA. Secondly, RT data tend to be positively skewed, so the Gaussian mean does not capture the full nature of the data (Ratcliff, 1979). Vincentile analysis can be used to plot vigilance decrements over time (Moss et al., 2016). Other models such as an ex-Gaussian distribution have been shown to better represent RT data (see Figure 2) (Leclaire et al., 2020; Schmiedek et al., 2007; Whelan, 2008). An ex-Gaussian model is a mathematical convolution of a Gaussian distribution and an exponential distribution and is characterized by three metrics: *mu* and *sigma* reflect the mean and standard deviation of the Gaussian component, respectively, and *tau* reflects the exponential component or 'slow tail' of the data (Schmiedek et al., 2007). The parameters

capture different features of the data, reflecting different cognitive processes within the task. *Tau* captures the exponential component, i.e., the long tail of the positive skew. This encompasses longer RTs, reflecting lapses in attention. Higher *tau* may therefore indicate poorer SA independent of baseline PS. *Mu* can be conceptualised as the speed of processing during the trials which did not involve a lapse in attention and *sigma* represents the intra-individual variation of processing speed during these trials. Ex-Gaussian analysis can therefore produce parameters that may reflect functioning in distinct cognitive domains from one neuropsychological task. These parameters have been used to reveal increased intra-individual response variability in clinical populations such as BD, MDD, and dementia (Gallagher, Nilsson, et al., 2015; Schumacher et al., 2019) and therefore offer a novel tool to measure SA more precisely. However, no other research to the author's knowledge has utilised ex-Gaussian modelling to test cognitive functions in mood disorders.

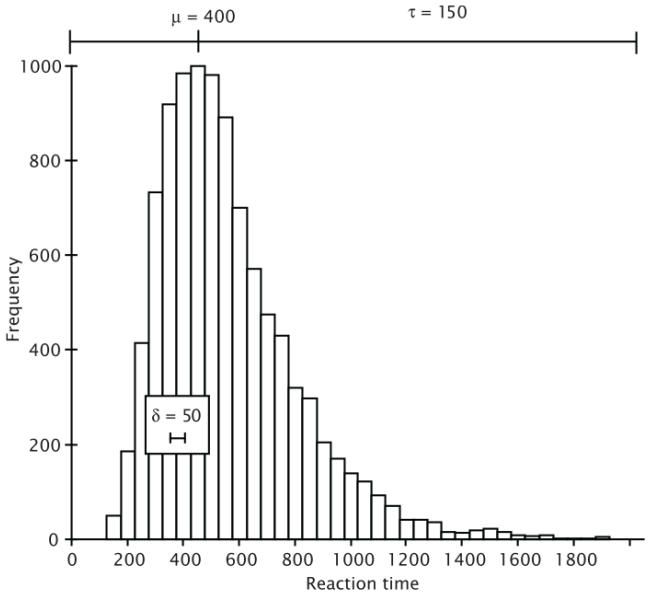

Figure 2. An example of an ex-Gaussian distribution of reaction times, including the parameters mu (μ), sigma (σ), and tau (τ). Image taken from Whelan (2008).

2.1.4. Dimension reduction and multivariate approaches

Testing multiple cognitive functions also poses problems for analysis: the use of univariate models to measure cognitive impairments in several functions requires corrections for multiple comparisons, which may limit results (Hu et al., 2022). Some studies attempt to mitigate this by creating composite scores calculated using several tests to represent cognitive domains, however the task impurity problem remains here, where the composite scores likely capture a range of cognitive abilities (Evans et al., 2014; Langenecker et al., 2010). Dimension

reduction techniques such as PCA and Factor Analysis can reveal latent factors in neuropsychological data, which may elucidate how cognitive functions are interrelated in clinical groups and how this profile may differ from healthy models (Gallagher et al., 2014). Another issue with a univariate approach is that it does not account for relationships between cognitive functions. In line with the idea of a cognitive hierarchy (Antila et al., 2011; Zaremba et al., 2019), cognitive functions should be considered together to account for relationships and overlap between them. Multivariate techniques should be utilised when testing associations between cognitive functioning and sets of other variables, e.g., clinical characteristics, and can reveal specific cognitive features that play a role whilst accounting for relationships between cognitive functions (Sparding et al., 2015).

2.2. Methodology for assessing brain-cognition associations

Similarly, a challenge with testing brain-cognition associations is dealing with multivariate datasets and assessing multiple relationships between several cognitive functions and brain metrics at once. A traditional approach in the mood disorders literature is to use linear models to test the association of one cognitive variable with structural features of several brain regions (Abé et al., 2018; Krabbendam et al., 2000; Sax et al., 1999). Other approaches involve splitting patient groups into subgroups based on cognitive performance (Macoveanu et al., 2021). These methods require corrections for multiple comparisons to reduce type-I errors, however this simultaneously increases type-II errors, arguably causing real associations to be missed (Rothman, 1990). Some studies reduce the number of comparisons by focussing on a specific region that theoretically may be of interest (Hanford et al., 2016; Sax et al., 1999; Yu et al., 2021). However, this approach may miss important relationships between cognitive functions and other areas of the brain. More complex models are therefore needed to detect brain-cognition associations that can simultaneously accommodate the multivariate cognitive and brain morphological data.

Given that the literature points towards the existence of complex associations between cognitive domains and brain structure in BD, multivariate techniques are needed to model such brain-cognition relationships (Genon et al., 2022; A. R. McIntosh & Mišić, 2013). Such techniques are starting to be used for these types of investigations, including machine learning techniques, canonical correlation analysis, and partial least squares analysis, however, this is an emerging field with few studies utilising these methods (Kebets et al., 2019; Mihalik et al.,

2022). The advantage of these techniques is their ability to limit the probability of committing type-I error by allowing for simultaneous comparisons among the variables without repeated tests. They also better reflect the complex, multimodal nature of behaviour and physiology (Sherry & Henson, 2005).

2.2.1. Canonical correlation analysis (CCA)

One promising method to test brain-behaviour associations is canonical correlation analysis (CCA). CCA is a data-driven multivariate approach that estimates the association between two sets of variables (*U* and *V*) by leveraging dimension reduction techniques (H. T. Wang et al., 2020). While dimension reduction techniques such as PCA attempt to reduce the number of variables in one set to components that capture variation in the data (via eigen decomposition of the correlation matrix), CCA reduces the number of variables in each set to latent variables (canonical variates) that maximise the correlation between two sets of data (via eigen decomposition of the *cross*-correlation matrix). The output of CCA is therefore correlated pairs of latent variates, where each pair is independent. The canonical variates are ordered with the first variate from each dataset (*U*1 and *V*1) explaining the largest proportion of covariance between the two sets. Canonical loadings represent the relationship between an original variable in the dataset and the canonical variate of its own dataset. Canonical *cross*-loadings represent the relationship between an original variable in the dataset and the canonical variate of the other dataset, allowing an estimation of which specific variables have the strongest association to the other dataset.

CCA therefore provides a method for testing associations between two multivariate datasets, without requiring adjustments for multiple comparisons. CCA is useful when there are high intercorrelations within variable sets (Lambert et al., 1988), which makes it appropriate for testing brain-cognition associations, since we might expect both cognitive variables and the brain structural variables to be interrelated within their own datasets. A limitation with CCA is the required sample size: there should be more subjects than the number of variables in both datasets. If this criterion is not met, data reduction techniques, such as PCA, can be employed to reduce the number of variables within each dataset before applying CCA (A. R. McIntosh & Mišić, 2013; H. T. Wang et al., 2020).

CCA is not new to the field of neuroimaging (Friston et al., 1995), however its application for brain-behaviour associations appears to be gaining popularity in more recent years. For

example, in healthy controls, CCA has been used to relate brain connectivity to lifestyle, demographic and behavioural data in large datasets such as the Human Connectome Project data (Smith et al., 2015). Another study used CCA to assess multivariate associations between brain structure and clinical, behavioural and cognitive variable in the Adolescent Brain and Cognitive Development (ABCD) data (Modabbernia et al., 2021). CCA has also been utilised to probe multi-domain brain-cognition associations in clinical populations, such as in groups of patients with temporal lobe epilepsy (Rodríguez-Cruces et al., 2020). In this study, the authors assessed structural brain metrics and performance on a neuropsychological test battery using CCA and followed this with a permutation test to assess the robustness of the results, which showed associations between structural brain metrics and cognitive performance.

There are limited studies to date that have used multivariate models to test brain-cognition associations in mood disorder groups. Rodrigue et al. (2018) used CCA to investigate associations of cognitive functioning across fourteen neuropsychological tests with brain morphology in patients with BD, schizophrenia spectrum disorders, and HCs. They grouped patients together to create an affective psychosis continuum and ran separate CCA models with each morphological metric (cortical volume, cortical thickness, surface area, and local gyrification indices). The results showed that the first canonical pair from each of the four analyses was associated with measures of general cognitive ability and larger volumes, thicker cortex, and smaller area in mostly frontal and parietal regions. They also interpreted the second canonical pair, which indicated associations of better WM with larger cortical volume in frontal and temporal regions, as well as slower RTs with thinner cortex in lateral frontal and temporal regions.

Other studies have used CCA on fMRI, clinical, and cognitive data and found that functional connectivity in frontostriatal and orbitofrontal areas was correlated with anhedonia and psychomotor slowing in treatment-resistant MDD (Drysdale et al., 2017). However, this study only used a single significance test (Wilk's lambda) to test the canonical correlation, and further analysis should arguably be done to test the robustness of this result, since CCA is prone to overfitting (Dinga et al., 2019). Dinga et al. replicated this analysis but added a permutation test and 10-fold cross-validation, where nine subsets were used as a training set and the remaining subset used as a test set. Strong canonical correlations were found for the first two canonical pairs, but these were not significant after permutation test and cross-validation. The sample differed between these two studies, as Dinga et al. included subjects

with MDD and anxiety disorders recruited from the general population with a wider range of symptom severity, ranging from sub-threshold depression in remitted MDD patients to severe symptoms. However, the results highlight the need to validate and replicate results.

Ang et al. (2020) investigated whether cognitive impairments were related to regional abnormalities in brain volume in children with depression, using data from the ABCD dataset (*n*=4,626 children aged 9-10 years old). They extracted the volume of 68 cortical regions of interest (ROIs) and 14 subcortical ROIs and regressed out intracranial volume and gender. Factor analysis was performed on eleven cognitive variables, which revealed three latent variables: language and reasoning, cognitive flexibility, and memory recall. CCA tested associations between the latent cognitive variables and cortical volume data: all three canonical variates were significant according to Wilk's lambda, however permutation tests and 10-fold cross-validation suggested that only the first canonical correlation was reliable. The first canonical correlation was associated with language and reasoning and showed the highest loadings with the volume of middle temporal gyrus, pars orbitalis, superior frontal gyrus, and superior parietal cortex areas. Together, these studies demonstrate that CCA, along with further tests of robustness, can be used to test brain-cognition associations.

2.3. SUMMARY AND AIMS OF THE THESIS

Processing speed, sustained attention and executive function appear to be impaired in BD and MDD, however, the exact nature of the impairment in PS and SA in mood disorders has not yet been established, likely due to heterogeneity in the methodology used in the literature. PS, SA, and EF may have a role as core cognitive processes that lead to secondary cognitive impairments, but few studies have accounted for interrelationships between cognitive functions or investigated the presence of a hierarchy of cognitive dysfunction in mood disorders. Whether core cognitive dysfunction is related to brain structure in patients is also unclear and most research to date is limited by univariate analysis. Combining multivariate techniques with methods that can more precisely fractionate distinct cognitive components could drive forward our ability to define cognitive phenotypes and test the role of core cognitive impairment in wider cognitive functioning and brain abnormalities in BD.

This thesis therefore had several aims: firstly, we aimed to establish the nature of impairments in PS and SA in people with BD and MDD by conducting a systematic review and meta-analysis

of the literature. In the review, we aimed to quantify the magnitude of impairment associated with each neuropsychological test (Chapter 3). Secondly, we sought to examine the role of PS, SA, and EF in wider cognitive functioning in people with BD and MDD using hierarchical regression analysis (Chapter 4). Finally, we aimed to investigate associations between abnormal brain structure and core cognitive dysfunction in people with BD using multivariate CCA (Chapter 5).

Chapter 3 SYSTEMATIC REVIEW AND META-ANALYSIS OF CORE COGNITIVE FUNCTIONS IN MOOD DISORDERS

3.1. INTRODUCTION

3.1.1. Background and objectives

Chapter 1 argued that impairments in core cognitive functions such as processing speed (PS) and sustained attention (SA) may play a role in wider cognitive dysfunction in mood disorders and that this should be formally tested. However, the nature of core cognitive impairments in BD and MDD should be established before their role in wider cognitive dysfunction is assessed. The literature has not yet reached a consensus on the presence and magnitude of impairments of PS and SA in mood disorders. Contradictory findings may be due to variations in the sample tested, including the current mood state of participants, and variations in neuropsychological methodology used to measure PS and SA. A more thorough, systematic investigation of the literature is warranted to assess the extent of impairment of PS and SA in BD and MDD. To our knowledge, no reviews of the literature have focussed on PS and SA in BD or MDD in the same way that EF has been more thoroughly reviewed and tested (Cotrena et al., 2020). Several reviews have investigated a similar question by meta-analysing studies investigating cognitive dysfunction more generally in mood disorders, including measures of PS and/or SA.

3.1.2. Systematic reviews of cognitive impairment in BD

Robinson et al. (2006) conducted the first systematic review and meta-analysis of 26 studies of cognitive function of euthymic BD patients compared to HCs. They found widespread impairment in euthymic BD with every measure showing a significant group difference. Tests of PS and SA showed medium ESs (CPT RT *d*=.60, CPT sensitivity *d*=.48, DSST *d*=.59, and TMT-A *d*=.52). Subsequent meta-analyses generally support these early fundings, reporting a similar pattern of moderate to large ESs for cognitive domains, including PS and SA (Arts et al., 2008; Bora et al., 2009; Kurtz & Gerraty, 2009; Lee et al., 2014; Mann-Wrobel et al., 2011; Torres et al., 2007). More recently, Cardenas et al. (2016) reviewed 51 articles and reported that euthymic BD show impairments in PS, attention, verbal fluency and verbal learning and memory. However, they did not meta-analyse data. Cullen et al. (2016) meta-analysed data from 15 studies to investigate the prevalence of cognitive impairment in euthymic BD. Fewer studies were included here perhaps due to stricter inclusion criteria such as only including

euthymic outpatient samples and studies with specific recruitment protocols. Cognitive deficits were present for a substantial proportion of euthymic BD: up to 57% showed impairment when impairment was defined at the 5th percentile impairment threshold. Both reviews noted that the results were inconsistent and were limited by the heterogeneity in neuropsychological tests and outcomes measures reported.

Bo et al. (2017) meta-analysed neuropsychological performance of BD patients in seven cognitive domains, including PS and attention/vigilance. BD performed worse than controls on all domains, with large ESs in PS and overall cognition; attention showed a medium ES. However, only 7 studies were included in their meta-analysis as the review was limited to studies that used the MCCB (Nuechterlein et al., 2008). Given the heterogeneity in the neuropsychological instruments used in the literature (Cardenas et al., 2016), investigating performance on specific neuropsychological tests may help to establish which tests are sensitive to core cognitive impairments. Bourne et al. (2013) meta-analysed performance euthymic BD samples on data from four tasks that were commonly used in the literature at the time: verbal learning tests, TMT, digit span, and the Wisconsin Card Sorting Test. They found impairments in many cognitive functions, including PS and attention/WM, even after controlling for age, IQ, and sex. The authors also found that residual mood symptoms appeared to confound their result, indicating that even though cognitive impairment is found in euthymia, it is important to investigate the cognitive profile of different mood states.

Tsitsipa and Fountoulakis (2015) reviewed 250 studies of cognitive performance in BD in manic, depressed, and euthymic states, and found impairments in almost every cognitive domain with medium ESs. They found a greater magnitude of impairment in symptomatic states compared to euthymia. Kurtz and Gerraty (2009) meta-analysed neuropsychological performance in BD and formally compared overall ESs of euthymic and symptomatic patients. They found a general impairment in BD patients compared to controls on most neuropsychological tests, but only verbal learning and verbal fluency tests showed a significant difference when comparing the impairment of euthymic and symptomatic BD groups. This suggests that cognitive performance is impaired in BD and may be worse in symptomatic states, but only for some neuropsychological tests.

3.1.3. Systematic reviews of cognitive impairment in MDD

A systematic review and meta-analysis by Lim et al. (2013) found that MDD showed poorer performance than controls on tests of attention, PS, as well as some tests of EF and immediate verbal memory. However, there were limited studies for some comparisons (e.g., only two studies with CPT data), perhaps because only PubMed and Cochrane databases were searched. The authors only reviewed studies that contained a sample with moderate or severe depression, so findings from people with mild depression were missed. Further, they analysed the cognitive domain of attention by grouping CPT scores with digit span and did not test SA specifically. The results suggest that not all neuropsychological instruments are sensitive to cognitive impairment in MDD, highlighting the need to assess performance on individual tests.

Parkinson et al. (2020) reviewed performance of symptomatic MDD and controls on specific cognitive tasks from 27 studies, however they restricted this to non-computerised instruments. MDD performed worse than controls on all 16 cognitive measures they tested with moderate to large ESs. The cognitive measures included Wechsler Digit Symbol Coding, TMT-A, phonemic fluency, and semantic fluency tests, which all showed significant differences, but all had heterogenous results. Other studies have focussed on cognitive functioning in the remitted phase, and generally show that remitted MDD show poorer cognitive functioning than controls on most cognitive domains, including PS and attention (Bora et al., 2013; Semkovska et al., 2019). Task-specific meta-analyses in these studies found that impairments in remission can be detected by TMT-A, verbal fluency tasks, RT tasks, and DSST. They also reported impairments in some tests of attention but not for SA specifically (Bora et al., 2013; Semkovska et al., 2019).

Other reviews have summarised the literature on cognitive functioning in both the symptomatic and remitted state: Ahern and Semkovska (2017) meta-analysed data from 31 studies of cognitive performance of first-episode MDD patients during a depressive episode and in remission. They reported impairments across most domains that varied from small to large ESs in acute phases, including tests of PS and SA. Remission was associated with normalisation in PS, but there were not enough data to analyse SA. One review meta-analysed data from specific tests from the Cambridge Neuropsychological Test Automated Battery (CANTAB; Cambridge Cognition, 2015) in both symptomatic and remitted MDD (Rock et al., 2014): currently depressed patients showed impairments in EF, attention (CPT), and memory, but not RT. Remitted patients showed deficits in EF and attention, but not memory, and there

were not enough data to meta-analyse RT. A more recent systematic review of cognitive impairment in MDD in both the acute and remitted state (Kriesche et al., 2022) found strong support for cognitive impairments in PS and SA in the acute phase of MDD. In the remitted phase, impairments in attention were found but to a lesser degree than the acute phase, whereas most studies did not find a decrease in PS for remitted MDD. However, studies were lacking for some cognitive functions, and they did not meta-analyse results.

3.1.4. Core cognitive functioning in mood disorders

Systematic reviews that specifically reviewed performance of PS or SA are scarce. Only one review focussed on attention in BD (Camelo et al., 2013): 110 studies were reviewed that included tests such as CPT, Stroop test, TMT-A, DSST, and go/no go tasks. They reported that BD showed impairments in attention that were present in euthymia and worsened in symptomatic states, however, data were not meta-analysed. One review focused on attention in MDD and meta-analysed performance in different sub-domains of attention (X. Wang et al., 2020). They found a significant overall deficit in global attention and in sub-domains of attention, including SA and psychomotor speed/attention, with moderate ESs. They meta-analysed individual test scores and found that MDD scored worse than controls on TMT-A, DSST, and CPT (RT and errors). One review meta-analysed tests of PS in MDD and found that depressed groups had slower RTs than controls, with a similar degree of slowing across tasks (White et al., 1997). However, this study included cognitively demanding tasks that we would not consider measures of PS, e.g., Stroop interference and face discrimination tasks. To our knowledge, there are currently no other systematic reviews of the impairment of, and/or methods used to test, core cognitive functions such as PS or SA in adults with mood disorders.

3.1.5. Summary and open questions in the literature

The systematic reviews and meta-analyses conducted so far generally indicate that BD and MDD perform worse than controls on most cognitive domains, including PS and SA, but that there is heterogeneity in cognitive functioning in mood disorders (Bo et al., 2017; Bourne et al., 2013; Cullen et al., 2016). Contradictory results in the literature may have resulted from heterogeneous methodology, including differences in neuropsychological tests used to measure PS and SA (Bortolato et al., 2015; Cardenas et al., 2016; Cullen et al., 2016; Parkinson et al., 2020). It would therefore be useful to study which instruments are associated with dysfunction in BD and MDD. Most reviews have focussed on cognitive domain scores rather

than individual cognitive tests (Bo et al., 2017; Tsitsipa & Fountoulakis, 2015) and some limit neuropsychological assessment to one assessment battery (Bo et al., 2017; Rock et al., 2014), a few pre-selected cognitive tests (Bourne et al., 2013), or only computerised tests (Parkinson et al., 2020). The reviews that investigated individual test scores require updating (Kurtz & Gerraty, 2009; Torres et al., 2007). Few studies have performed meta-analysis across mood states, with many restricting data to euthymic patients (Bourne et al., 2013; Cardenas et al., 2016; Cullen et al., 2016). Some studies did include all mood states but do not formally analyse differences between them (Bo et al., 2017). The literature tends to investigate MDD and BD patients separately. However, some studies have failed to find evidence of distinct cognitive profiles between MDD and BD, and instead note that cognitive impairments are more severe in symptomatic states regardless of diagnoses (Porter et al., 2015). It may therefore be important to investigate core cognitive functioning in both BD and MDD. An updated review is therefore necessary to investigate the nature of impairments in PS and SA in mood disorders by investigating performance across different neuropsychological tests and in different mood states in BD and MDD.

3.1.6. Objectives and scope of the review

This review aimed to systematically review the published literature on PS and SA performance in adults with MDD or BD. It aimed to assess the neuropsychological methods used in studies and to quantify, through a meta-analysis, the magnitude of impairments reported in BD and MDD compared to HCs for each standardised neuropsychological tests and outcome measure. Where possible, we aimed to separate the meta-analysis by mood state, to test differences between euthymic and symptomatic states.

3.2. METHODS

3.2.1. Scoping and protocol development

During the scoping process, a search for published systematic reviews or systematic review protocols that investigated PS or SA in adults with BD or MDD was conducted in February 2020 in the Cochrane Library, the PROSPERO library, and Web of Science. No published reviews or review protocols were found. Scoping searches were also conducted to assess the number of search results and develop our search strategy and search terms. A protocol for this systematic review and meta-analysis was pre-registered before the review began according to PRISMA-P

guidelines (Moher et al., 2015). The protocol was published on the International Prospective Register of Systematic Reviews (PROSPERA; CRD42020179443; Little & Gallagher (2020); https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42020179443).

3.2.2. Eligibility criteria

A detailed eligibility checklist of inclusion and exclusion criteria was used during screening, which is detailed in Table 18 in Appendix A. In sum, the following inclusion criteria were applied: original research published in peer-reviewed journals from 1994 onwards; English language; studies of adults aged 18-65 years old with a primary diagnosis of BD or MDD with no comorbidities; inclusion of a HC group; inclusion of at least one objective test of PS or SA with a numeric score. 1994 was used as this was the year in which the DSM-IV and ICD-10 diagnostic criteria were published. Brain imaging or eye-tracking studies were excluded to limit the number of search results and because these studies typically contain smaller samples.

3.2.3. Search

3.2.3.1. Databases

Research suggests that systematic reviews should search subject-specific databases as well as general databases to gain adequate coverage of the literature (Bramer et al., 2017). Bramer et al. specifically suggested that optimal searches should search at least Embase, Medline, Web of Science, and Google Scholar, with the option of adding other specialized databases (e.g., PsycINFO). Therefore, several electronic databases were searched: Medline, Embase, and APA PsychINFO (all accessed via Ovid) as subject-specific databases; and Web of Science as a general database. We did not include Google Scholar because our search yielded a large number of references using the subject-specific databases and Web of Science alone. Hand searches of reference lists of relevant articles were also conducted.

3.2.3.2. Search strategy

The search terms were developed in Ovid and then applied to Web of Science. An example of the complete final search strategy used in Ovid is presented in Table 19 in Appendix A. The search comprised of key word variants relating to (a) cognition or processing speed or sustained attention, and (b) major depressive disorder or bipolar disorder. The initial search terms used were: (Neurocog* OR neuropsychol* OR cogniti* OR attention* OR “sustained attention” OR vigilance OR “mental speed” OR “processing speed” OR “speed of processing”

OR “speed of information processing” OR psychomotor) AND (bipolar OR manic OR mania OR manic-depress* OR BD or depress* OR MDD OR “mood disorder*” OR “affective disorder*”). The initial search was restricted to original journal articles reported in English language published between January 1994 to the date of the search (April 2020).

Searching through titles, keyword headings, keyword heading words, and abstracts using the terms detailed above yielded a very large number of results (**294,565** in Ovid alone), so several steps were taken to reduce this to a manageable number. Since our search terms contained words that can be used within abstracts in other contexts that were not relevant here (e.g., “attention”), the search was restricted to only keywords and titles, which reduced the number of results (**32,963** in Ovid). A portion of the resulting references were then checked by eye and some frequently occurring irrelevant topics were identified; these were ADHD, dementia, and Cognitive Behavioural Therapy (CBT). Additional search terms were subsequently added to the search strategy to further narrow down the results. These were: “NOT ('attention deficit hyper*' OR ADHD OR ADD OR 'attention deficit disorder*') NOT (therapy OR CBT OR 'cognitive-behavio*') NOT ('dement*')”. Adding these terms further reduced the search results (**21,990** in Ovid). All additional restrictions that were available in Ovid to specify only original peer reviewed journal articles were then selected; these restrictions were specific to each database (see Table 19 in Appendix A for details). The same search strategy was applied to Web of Science in April 2020.

De-duplication of the Ovid search results was done in three stages: first, the de-duplication function in Ovid was used, followed by the Systematic Review Assistant-Deduplication Module software (SRA-DM; Rathbone et al., 2015). SRA-DM is an automated de-duplicator that has higher specificity (100%) and sensitivity (84%) than commonly used techniques such as Endnote (Rathbone et al., 2015). Finally, the remaining duplicates were identified using the automated duplicate detector in the Rayyan web app for systematic reviews (Ouzzani et al., 2016) and these duplicates were manually checked. The Ovid and Web of Science results were merged and de-duplicated by using SRA-DM followed by a manual check of the remaining duplicates using Rayyan. Duplicates were removed from the Web of Science search results and the remaining search results were uploaded to Rayyan for screening.

An updated search was conducted in October 2021 to ensure relevant and up to date references were captured. The same search strategy was used, with the only difference being

a limit on publication date of 2020 to the date of the search. Since the Web of Science search did not yield many additional references that met our criteria over and above the original Ovid search in April 2020 (*n*=9), only the specialised databases were searched at this stage (Medline, Embase, and APA PsychINFO).

3.2.3.3. Search verification

Given that post-hoc restrictions were added to our search criteria, the sensitivity of the search strategy was carefully tested and verified by checking whether a list of key papers that were known to the authors to be relevant (*n*=12) had been detected by the search. The references of relevant published reviews and retrieved articles were also hand-searched for further relevant articles not included in the initial search. All key papers were included in the search results; thus, we considered our search criteria adequate.

3.2.4. Screening and article selection

Two reviewers piloted the eligibility checklist independently using three references from the search results and the checklist was deemed appropriate for screening. The titles and abstracts of each reference were screened for relevance independently by two reviewers using Rayyan. Each reviewer was blind to the decisions of the other and independently recorded decisions and reasons for exclusion in Rayyan. When blinded screening was complete, Rayyan was unblinded and any discrepancies were discussed until consensus was reached. The full texts of all articles that survived title and abstract screening were retrieved from Google Scholar and uploaded to Rayyan. Two reviewers independently screened the full texts using the eligibility checklist of inclusion and exclusion criteria. Rayyan was used to record decisions and reasons for exclusion. The reviewers were blind to the decisions of the other reviewer throughout full-text screening. When both reviewers had completed screening, discrepancies between reviewers were resolved by consensus discussion.

If a paper met most of the inclusion criteria but had information or data missing, the corresponding authors were contacted via email to ask for the relevant information or data. Likewise, if studies otherwise met criteria but data for BD or MDD groups were mixed with another group (e.g., BD patients grouped together with schizophrenia patients), authors were contacted to acquire disaggregated data. Studies were excluded if there was no response from authors after 6 weeks, or if authors specified that the information or data were not available.

Care was taken to exclude multiple publications of data derived from the same sample or overlapping samples. Authors were contacted if it was unclear whether studies involved the same samples. In the case of redundant data, i.e., where the same dataset has been reported in two or more articles, the article reporting the largest sample was considered in the analysis. Several articles reporting the same sample were included if the articles reported different neuropsychological measures and/or outcome variables, or if they had tested patients at two distinct time points in different mood states.

3.2.5. Data extraction

Data were extracted using a standardised data extraction form in Excel. The form was piloted by two reviewers using five papers each. The form was then edited as appropriate to address any issues and allow for the full breadth of data to be extracted. Data collected included: citation details, paper characteristics, methods and materials, sample demographics, clinical characteristics, neuropsychological test information, and outcome variable data. Table 20 in Appendix A details of the types of data extracted. Data extraction was carried out by Reviewer 1 (BL). A second reviewer independently extracted data from a randomly chosen subset of papers (approximately 20%) to check for consistency. Data extracted by each reviewer on these papers were cross-checked, checking for accuracy and completeness; no discrepancies were found. Mean and standard deviation (SD) of neuropsychological test scores were extracted for meta-analysis. If studies reported median scores and range in place of the mean and SD, we contacted authors to ask for group means and SDs. If the authors did not respond, we assumed the data were skewed and did not use the data.

3.2.6. Data synthesis

Data were collated into tabulated summaries of study characteristics for each article. Data were meta-analysed for a given comparison if there were more than two studies with data available for that comparison. Data synthesis was grouped according to mood disorder type (BD or MDD) and neuropsychological outcome measures used. For example, simple-RT for BD patients and controls constituted one comparison. Meta-analysis was separated by outcome measure (e.g., RT, number correct, or errors for CPT) to allow for more detailed investigation of different methodologies. Where possible, data were sub-grouped based on mood state. If studies reported composite scores of PS or SA tasks and they were not able to provide disaggregated data for the individual tests, we meta-analysed composite scores that

contained appropriate neuropsychological tests. Some studies were excluded for if the composite scores contained neuropsychological tests that we did not consider measure of PS or SA (e.g., digit span). Where studies split patients into groups and provided data for separate groups (e.g., grouped patients based on medication type), data were combined into one group using the Cochrane formula (Higgins et al., 2022).

Meta-analysis was conducted by pooling neuropsychological data from each article for each comparison. Review Manager version 5.4 (The Cochrane Collaboration, 2014) was used to calculate the standardized mean differences (SMDs) and 95% confidence intervals (CIs) for each comparison between patient groups and HCs. SMD (i.e., Cohen's *d*) was used to measure ES, as opposed to mean difference, to allow us to pool data from studies that used different scoring methods. SMD were interpreted according to Cohen (1988; small 0.2-0.5, medium 0.5-0.8, and large >0.80). Forest plots were generated using Review Manager to illustrate the overall result and sub-group analysis.

We expected heterogeneity in samples and methodology across studies based on other meta-analyses of neuropsychological functions (Bo et al., 2017; Parkinson et al., 2020), therefore we did not assume a common ES for studies and random-effects models were used to pool ESs to avoid overestimation of the overall difference (Mikolajewicz & Komarova, 2019). Heterogeneity of weighted average ES between studies was assessed using statistics: *Q,* which indicates the presence/absence of heterogeneity using a critical value of .05; and I^2, which reflects the degree of heterogeneity, i.e., the percentage of variation across studies that is due to heterogeneity (Higgins et al., 2003). A value of 50% or higher was used to indicate meaningful heterogeneity between the studies: 0% indicates no heterogeneity, and 25%, 50%, and 75% indicate low, moderate and high heterogeneity, respectively (Higgins et al., 2003).

3.2.7. Sensitivity and assessment of risk bias

To explore the robustness of the meta-analysis conclusions, a sensitivity analysis was performed for each comparison, to see if removing each study individually changed the overall result. Where possible, we also tested whether removing each mood state subgroup changed the overall result.

Potential sources of bias in our systematic review methodology were discussed between reviewers and reported narratively. Risk of bias across studies was assessed using funnel plots and Egger's test of publication bias. Funnel plots are scatterplots that plot effect estimate (i.e.,

SMD; x-axis) against study size (i.e., standard error of the SMD; y-axis) for each study in the comparison. We expected larger (i.e., more precise) studies to have SMDs close to the centre on the x-axis, and smaller studies to have SMDs scattered symmetrically around the centre, creating an inverted funnel shape that can be used to assess the relationship between magnitude and precision of effect estimates. Asymmetry in the funnel shape can indicate small study bias, reflecting publication bias (Egger et al., 1997; Higgins, Thomas, et al., 2022). Egger's test formally tests funnel plot asymmetry by conducting a linear regression of the effect estimates and their standard errors, where a significant result (p<.05) indicates asymmetry (Egger et al., 1997). Funnel plots were generated using Review Manager and visually inspected. Egger's test was conducted using the *metafor* package in R version 4.1.2 (R Core Team, 2021; Viechtbauer, 2010). Funnel plots and Egger's tests were only conducted for comparisons with at least 10 studies, as per Cochrane recommendations (Higgins et al., 2022; Sterne et al., 2011), and interpreted to assess publication bias due to missing results. We also considered the risk of publication bias by categorising papers into three categories: those with a primary aim to assess impairments in PS and/or SA; those with a primary aim to assess cognitive functioning generally; and those that contained neuropsychological data but did not assess neuropsychological functioning as a primary aim.

3.3. Results

3.3.1. The evidence base

Figure 3 shows a flow diagram of the literature search and screening process. In total, **21,009** records were identified in the literature searches after de-duplication. The titles and abstracts of these records were screened and 18,583 were removed. The full texts of **2,426** records were sourced and assessed for eligibility. The corresponding author of 503 records was contacted to request more information and/or data. 2,323 records were excluded during full-text screening and after contacting authors. The main reasons for exclusion at full-text screening were: no HC sample (*n*=646), data missing/not available (*n*=465), no tests of PS or SA (*n*=418), not an original article (*n*=143), no MDD or BD group (*n*=141), comorbidity (*n*=132), redundant data (*n*=124), no neuropsychological measures (*n*=78), older adults or children (*n*=58), diagnosis tool (*n*=55), and imaging or eye tracking study (*n*=47). After full text screening, **103** studies were included in the review and **101** were eligible for meta-analysis.

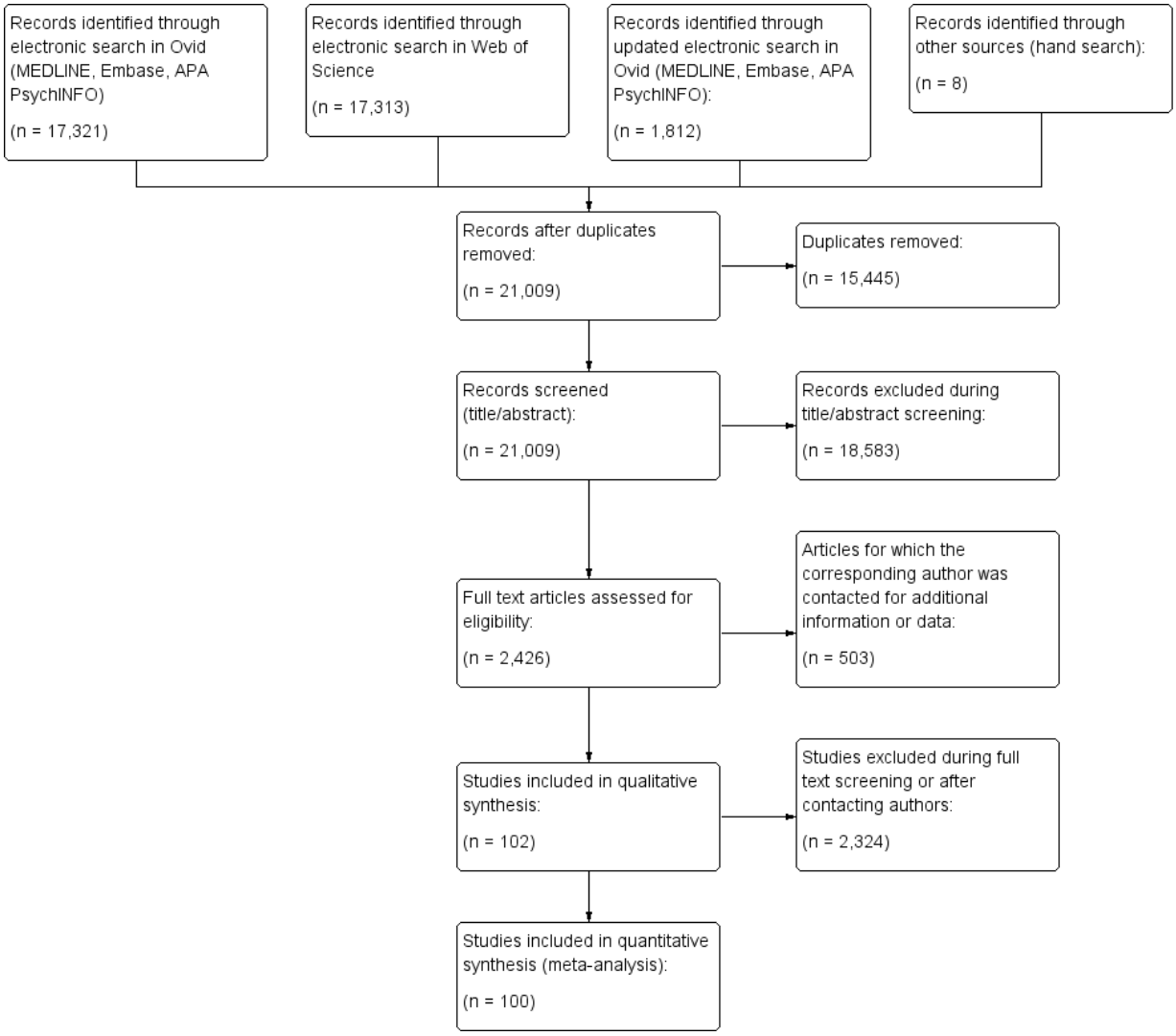

Figure 3 A flow chart of the systematic review literature search and screening process

3.3.2. Characteristics of studies

Given the large number of studies (*k*=103) included in this review, an overview of the paper characteristics is reported in the following text and summarised in Table 21 in Appendix A.

3.3.2.1. Participants

Fifty-two studies included a sample of patients with BD, 40 studies included a sample of MDD, and eleven studies included samples of both MDD and BD. A total of *n*=8,913 patients (*n*=3,452 BD and *n*=5,461 MDD) and *n*=8,016 HCs were included in the meta-analysis. The sample size of patients ranged from *n*=8 to *n*=766. The sample size of the HCs ranged from *n*=11 to *n*=668. Most studies used DSM-IV diagnostic criteria (*k*=83), with the remaining studies using ICD-10 (*k*=11), DSM-5 (*k*=7), and DSM-III (*k*=1) criteria.

3.3.2.2. Clinical characteristics

An overview of the clinical characteristics of the included studies is presented in Table 1. Fifty-one studies included a sample of outpatients, 11 studies included a sample of inpatients, 17 studies had a sample of both inpatients and outpatients, 3 studies recruited participants from the community, and 21 studies did not state the setting. Of the studies that included a BD group, *k*=22 only included BD-I, *k*=5 only included BD-II, *k*=15 included both BD-I and BD-II (one of which separated the groups), and *k*=21 did not specify the type of BD. *k*=37 included a sample of BD patients who were euthymic, *k*=10 included a sample of BD patients who were depressed, *k*=2 included a sample of BD patients who were manic, *k*=2 included a sample of patients who were symptomatic (manic and/or depressed), *k*=2 included a sample of BD patients that included various mood states including euthymia, and *k*=12 did not state the mood state of participants. Thirty-six studies included a sample of MDD patients who were depressed, *k*=10 included a sample of MDD patients who were remitted, and *k*=8 did not specify the mood state of participants.

Twenty-two studies included a sample of only non-psychotic participants (*k*=6 BD, *k*=14 MDD, *k*=2 BD and MDD), 7 studies included a sample of only participants with current psychosis or a history of psychosis (all were BD samples), 20 studies included a sample where a portion of the participants had psychosis or a history of psychosis (*k*=16 BD, *k*=4 BD, and MDD), and *k*=54 studies did not provide information on psychotic features of the sample (*k*=23 BD, *k*=26 MDD, *k*=5 BD, and MDD). Sixty-four studies included a sample of patients who were currently taking medication. Fourteen studies included a sample who were unmedicated at the time of testing and five studies included a drug naïve sample. Two studies grouped participants by the type of medication they were taking, and one study grouped participants into medicated and unmedicated groups. Seventeen studies did not state the medication status of participants.

3.3.2.3. Neuropsychological methodology

Figure 4 displays a heatmap of the number of studies that included each neuropsychological outcome measure for BD and MDD, split by mood state. Eighty-nine (87%) studies included at least one test of PS. Of these studies, the most common tests of PS were DSST and TMT-A, both reported in 48 studies (54%). Verbal fluency was included in 29 studies (33%). The remaining tests of PS we extracted from the studies were the Stroop task (simple trials, *k*=13), RT tasks (choice-RT *k*=10, and simple-RT *k*=5), and seven studies reported composite PS scores

from standardised test batteries, usually MCCB PS composite score (Nuechterlein et al., 2008). Note that composite scores were only extracted for studies that did not provide disaggregated data for the individual tests included in the composite score. There were two studies that used other tests of PS that were not meta-analysed; these were the Timed Chase Test and the Speed and Capacity of Language Processing Test (SCOLP) test. Eleven studies also included tests of motor speed; these were not analysed here. Note that not all these studies intended to use the neuropsychological tests reported here as a measure of PS and may have intended to measure other cognitive functions (e.g., verbal fluency is often used to measure EF).

Forty-one (40%) studies included at least one test of SA. Of these studies, the most common test for SA was the CPT, which was reported in 37 studies (90%). Five other tests of SA were reported; these were grouped together and included a composite SA score from the CNS Vital Signs battery (CNSVS; Gualtieri & Johnson, 2006; *k*=2; these studies did not provide disaggregated data from the composite score), the Attentional Network Test (ANT; *k*=1), ANTI-vigilance (*k*=1), and the Paced Auditory Serial Addition Test (PASAT; *k*=1). The MCCB attention/vigilance composite score only contains data from a CPT, so there was no need to acquire aggregate data.

Clinical characteristics		N studies (*k*)
Mood disorder diagnosis	Bipolar Disorder (BD)	52
	Major Depressive Disorder (MDD)	40
	Both BD and MDD	11
Diagnostic criteria	DSM-III	1
	DSM-IV	83
	DSM-5	8
	ICD-10	11
BD subtype	BD-I	22
	BD-II	5
	BD-I and BD-II (grouped together)	14
	BD-I and BD-II (separated)	1
	not stated	21
Setting	inpatients	11
	outpatients	51
	inpatients and outpatients	17
	community	3
	not stated	21
Mood state*	euthymic/remitted	47
	depressed	46
	manic	2
	various (grouped together)	4
	not stated	20
Medication status*	drug naïve	5
	unmedicated	15
	Medicated (various medications)	64
	Medicated (grouped based on medication)	2
	not stated	17
Other clinical characteristics*	whole sample with psychotic features	7
	mixed sample with and without psychotic features	20
	no psychosis	22
	first episode	6
	mixed sample first episode and multiple episodes	5
	not stated	57

*Table 1: Summary of clinical characteristics of studies. *Totals do not equal the number of studies, as some studies contained more than one sample (disaggregated data) and/or not all studies reported this information. BD=bipolar disorder; MDD=major depressive disorder; DSM=Diagnostic and Statistical Manual of Mental Disorders; ICD=International Classification of Diseases.*

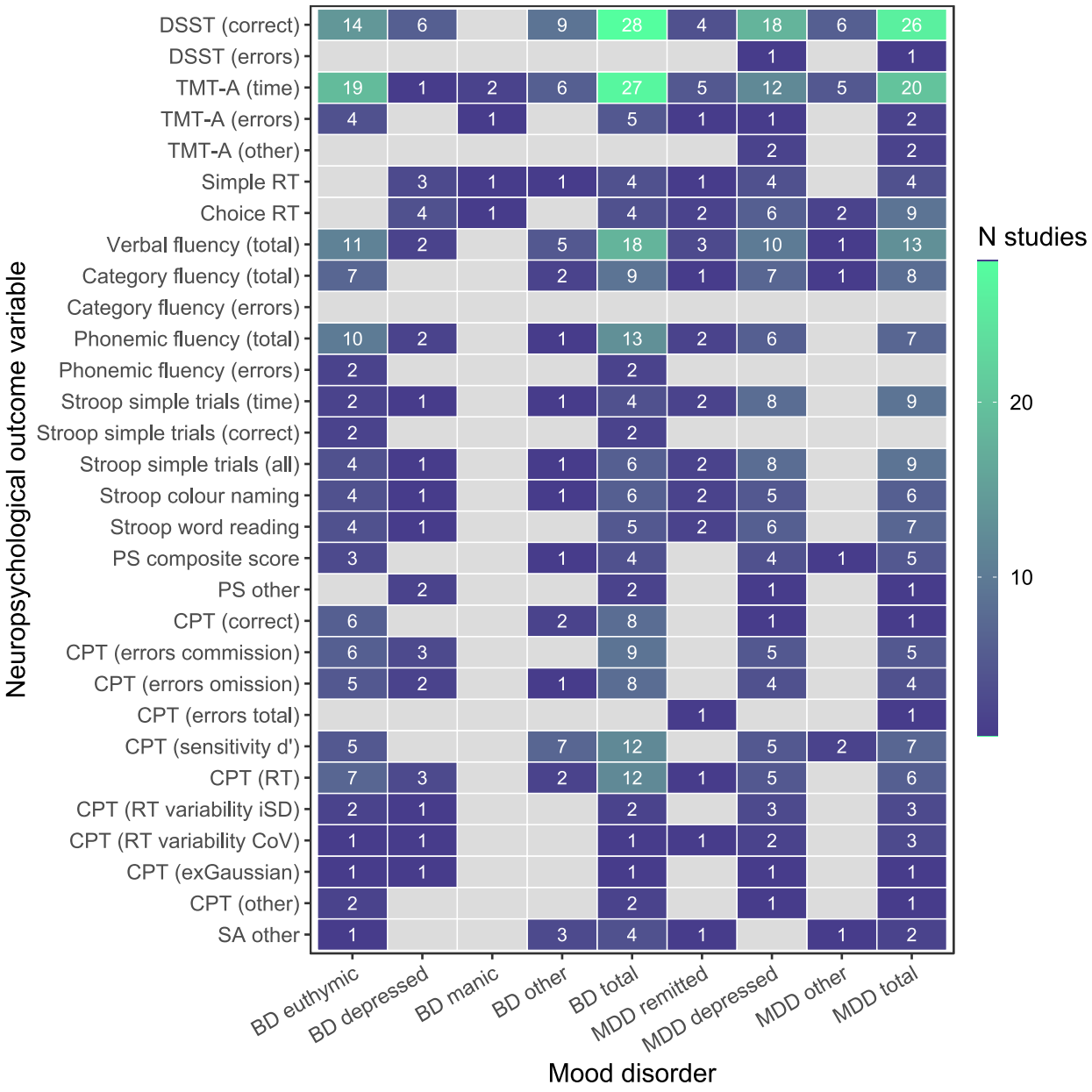

Figure 4: A heatmap of the number of studies that included each neuropsychological test for BD and MDD in each mood state. A grey cell indicates that no studies were found for that outcome measure and sample. BD=bipolar disorder; MDD=major depressive disorder; DSST=Digit Symbol Substitution Test; TMT=Trail Making Test; RT=reaction time, PS=processing speed; CPT=Continuous Performance Test; iSD=individual standard deviation; CoV; coefficient of variation; SA=sustained attention.

	Bipolar Disorder								Major Depressive Disorder							
	k	***ks***	**N BD**	**N HC**	**ES**	**CI 95%**	***Z***	***p***	***k***	***ks***	**N MDD**	**N HC**	**ES**	**CI 95%**	***Z***	***p***
Processing speed																
DSST (correct)	28	29	1791	2287	-0.927	-1.115 to -0.740	9.699	<.001	26	28	2767	3017	-0.704	-0.870 to -0.538	8.297	<.001
TMT-A (time)	27	28	1597	1416	0.576	0.382 to 0.771	5.808	<.001	20	22	2739	2579	0.569	0.391 to 0.747	6.272	<.001
TMT-A (errors)	5	5	244	170	0.699	-0.041 to 1.439	1.851	.064								
Simple-RT	4	5	84	194	0.636	0.362 to 0.910	4.546	<.001	4	5	308	269	0.315	-0.018 to 0.648	1.853	.064
Choice-RT	4	5	147	217	0.646	0.313 to 0.980	3.797	<.001	9	10	1277	876	0.289	0.137 to 0.440	3.737	<.001
Verbal fluency (correct)	18	18	873	851	-0.545	-0.722 to -0.368	6.048	<.001	13	14	1167	1083	-0.470	-0.602 to -0.339	7.003	<.001
Category fluency	9	9	556	361	-0.609	-0.858 to -0.360	4.799	<.001	8	9	990	916	-0.476	-0.641 to -0.310	5.643	<.001
Phonemic fluency	13	13	707	591	-0.610	-0.786 to -0.434	6.794	<.001	7	8	351	317	-0.506	-0.769 to -0.243	3.767	<.001
Stroop simple trials	6	6	282	445	-0.785	-0.953 to -0.618	9.183	<.001	9	10	1029	929	0.725	0.391 to 1.059	4.256	<.001
Stroop colour naming	6	6	282	445	-0.895	-1.064 to -0.725	10.350	<.001	6	7	875	788	0.593	0.330 to 0.855	4.420	<.001
Stroop word reading	5	5	199	177	-0.874	-1.100 to -0.648	7.583	<.001	7	8	905	845	0.561	0.336 to 0.786	4.882	<.001
PS composite	4	4	348	943	-0.955	-1.212 to -0.699	7.298	<.001	5	5	653	935	-0.862	-1.131 to -0.544	6.299	<.001
Sustained attention																
CPT correct	8	8	296	296	-0.446	-0.673 to -0.220	3.865	<.001								
CPT errors of commission	9	9	361	341	0.453	0.147 to 0.758	2.902	.004	5	5	185	179	0.767	0.552 to 0.981	7.015	<.001
CPT errors of omission	8	8	370	364	0.473	0.267 to 0.679	4.495	<.001	4	4	155	149	0.766	0.532 to 1.001	6.409	<.001
CPT sensitivity (d')	12	12	718	781	-0.645	-0.840 to -0.451	6.513	<.001	7	7	751	632	-0.403	-0.587 to -0.218	4.274	<.001
CPT RT	12	12	511	513	0.187	-0.178 to 0.552	1.004	.315	6	6	228	222	0.597	0.011 to 1.182	1.997	.046
CPT RT variability (iSD)	2	3	153	153	0.393	-0.518 to 1.304	0.845	.398	3	3	100	104	0.792	0.370 to 1.214	3.682	<.001
CPT RT variability (CoV)									3	3	91	97	0.669	-0.006 to 1.344	1.944	.052
SA other	4	4	229	778	-0.642	-1.131 to -0.153	2.572	.010								

Table 2: Summary of meta-analysis results for each comparison. k=number of studies; ks=number of samples; BD=bipolar disorder; MDD=major depressive disorder; HC=healthy control; ES=effect size; DSST=Digit Symbol Substitution Test; TMT=Trail Making Test; RT=reaction time, PS=processing speed; CPT=Continuous Performance Test; iSD=individual standard deviation; CoV; coefficient of variation; SA=sustained attention. k *and* ks *may differ where one or more studies had two separate samples, e.g., one euthymic BD group and one depressed BD groups. The effect size reported is the standardised mean difference (Cohen's d). Negative ESs reflect a test where higher score equals better performance.*

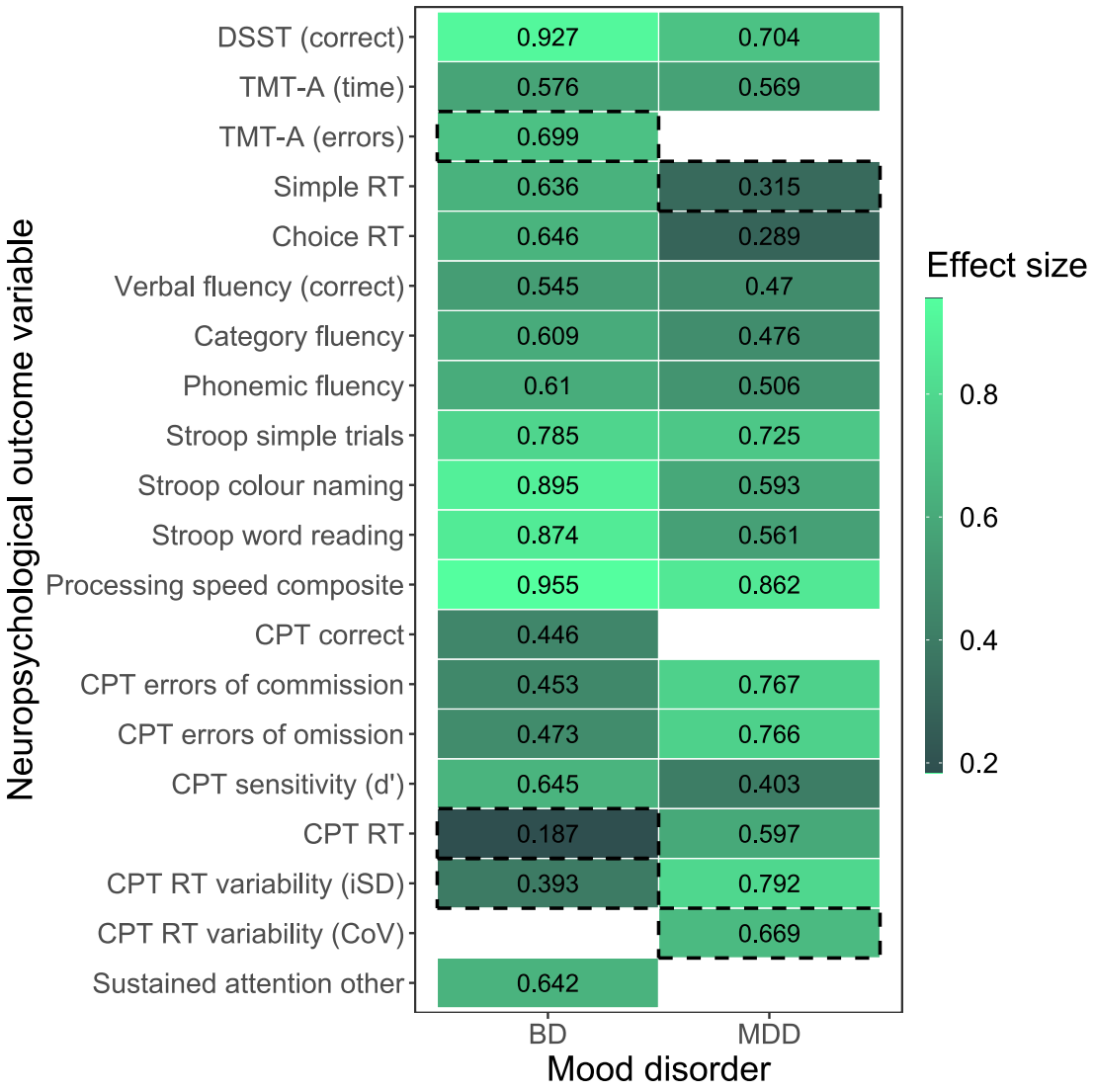

Figure 5: Summary of overall effect sizes (Cohen's d*) on each neuropsychological outcome variable for each diagnostic group, pooled across mood states. Blank white cells indicate where there were not enough data to perform meta-analysis. Dashed black outlines around cells indicate that the overall effect size was not significant (p>.05). DSST=Digit Symbol Substitution Test; TMT=Trail Making Test; RT=reaction time; CPT=Continuous Performance Test; iSD=individual standard deviation; CoV; coefficient of variation.*

3.3.3. Summary of meta-analysis

The overall results of the meta-analysis are summarised in Table 2. Figure 5 also illustrates the overall effect sizes for each neuropsychological outcome variable for BD and MDD. Forest plots are presented to illustrate the results for some individual neuropsychological tests and outcome measures, including subgroup analyses. BD performed worse than HCs on most measures of PS and SA, except number of errors on TMT-A, CPT RT, and CPT RT individual standard deviation (iSD). Large ESs were found for DSST (*d*=0.93), Stroop colour naming (*d*=0.89) and word reading (*d*=0.87), and PS composite scores (*d*=0.95). Most other significant comparisons were moderate ESs, ranging from 0.55 to 0.78; small ES were found for several

CPT outcome variables: number correct, and errors of commission and omission (ranging from 0.45-0.47). MDD performed worse than HCs on most measures of PS and SA, except simple-RT and CPT coefficient of variation (CoV). A large ES was found for PS composite scores (*d*=0.86); several CPT scores were nearing large ESs (errors of commission and omission, and RT iSD; all > 0.75). Most other significant comparisons were moderate ESs, ranging from 0.51 to 0.79. Small ESs were found for choice-RT, verbal fluency, category fluency, and CPT d', ranging from 0.29-0.48.

3.3.4. Meta-analysis of bipolar disorder

3.3.4.1. DSST

The number of correct responses on the DSST was significantly lower for BD patients than controls, as shown in Figure 6. For this comparison, there were enough studies to group samples based on mood state, including euthymic (*k*=14) and depressed groups (*k*=6), as well as a group of studies that did not specify the mood state of participants or contained mixed samples (*k*=9). The overall group difference was significant in each of the subgroups (all *p*s<.05) and there was no significant difference between subgroups (*p*=.58). There was significant heterogeneity overall, as well as within subgroups. Only one study showed the opposite effect, where patients performed better than controls (W. Zhang et al., 2020). Sensitivity analysis showed that removing this study did not change the overall outcome or the individual subgroup outcomes. There were no studies that used other metrics from the DSST, e.g., time to complete or number of errors, when testing BD patients.

3.3.4.2. TMT-A

Most studies that used TMT-A reported the time taken to complete the task as the outcome measure. BD took longer to complete TMT-A than HCs, as shown in Figure 7. Studies were grouped based on mood state, including euthymic (*k*=19), manic (*k*=2), and unspecified/mixed samples (*k*=6); only one study reported data for a depressed sample. The overall group difference was significant in each of the subgroups (all *p*s<.05), except the unspecified/mixed sample (*p*=.82). There was a significant difference between subgroups (*p*<.001), which appeared to be driven by the manic subgroup; when this subgroup was removed, there was no longer a significant difference between groups. Conducting subgroup analysis for pairs of subgroups showed that the manic and unspecified/mixed subgroups were significantly different from the euthymic and depressed subgroups, and each other. There was significant

heterogeneity overall, as well as within the euthymic and unspecified/mixed subgroups. Two studies with unspecified/mixed samples found that BD performed better than HCs (Lewandowski et al., 2016, 2020).

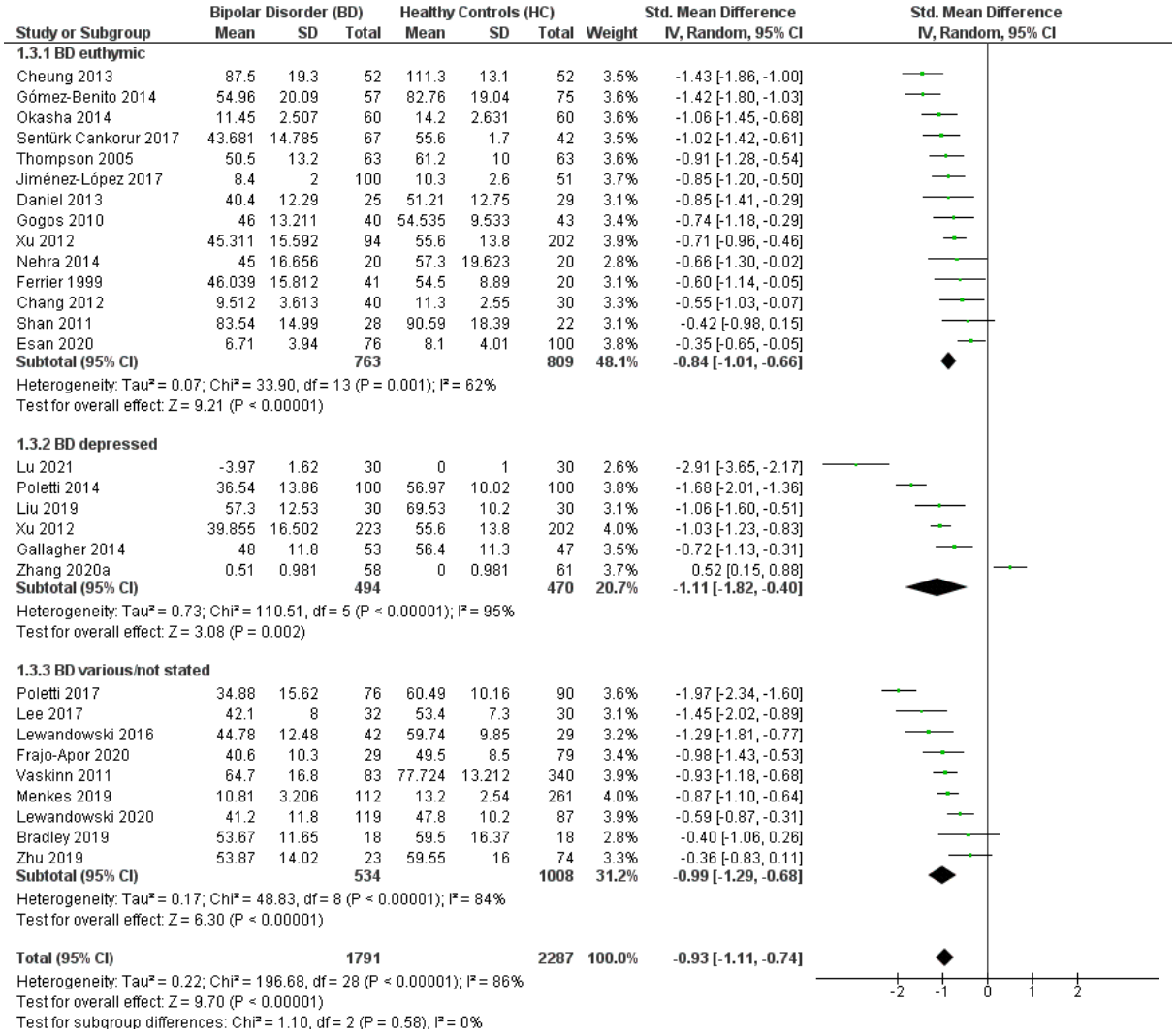

Figure 6: Forest plot of the comparison between BD patients and healthy controls on DSST number correct scores.

Fewer studies (*k*=5) used number of errors as an outcome measure for TMT-A, see Figure 8; four of these studies included a sample of euthymic patients, and one included a sample of manic patients. The ES appeared to be moderate (*d*=0.70), however, there was no overall group difference (*p*=.06). This may have been due to significant heterogeneity overall; only two of the studies showed significant group differences.

3.3.4.3. Reaction time

Performance on simple-RT tasks was significantly worse (i.e., longer RTs) for BD patients than controls (*p*<.05) with no significant heterogeneity, as shown in Figure 9. Three studies included

samples of depressed BD patients, one study included a sample of manic patients, and one study had an unspecified/mixed sample. A sensitivity analysis showed that removing the manic and/or unspecified samples did not change the overall effect.

Overall, BD patients had longer RTs on choice-RT tasks than HCs, with a moderate ES (*d*=0.65), see Figure 10. Four of the five samples were depressed patients and one sample was patients who were manic (note that two samples came from the same paper; Sweeney et al. (2000) included both a depressed sample and manic sample). A sensitivity analysis showed that removing the manic sample did not change the overall effect. There was no significant heterogeneity across studies.

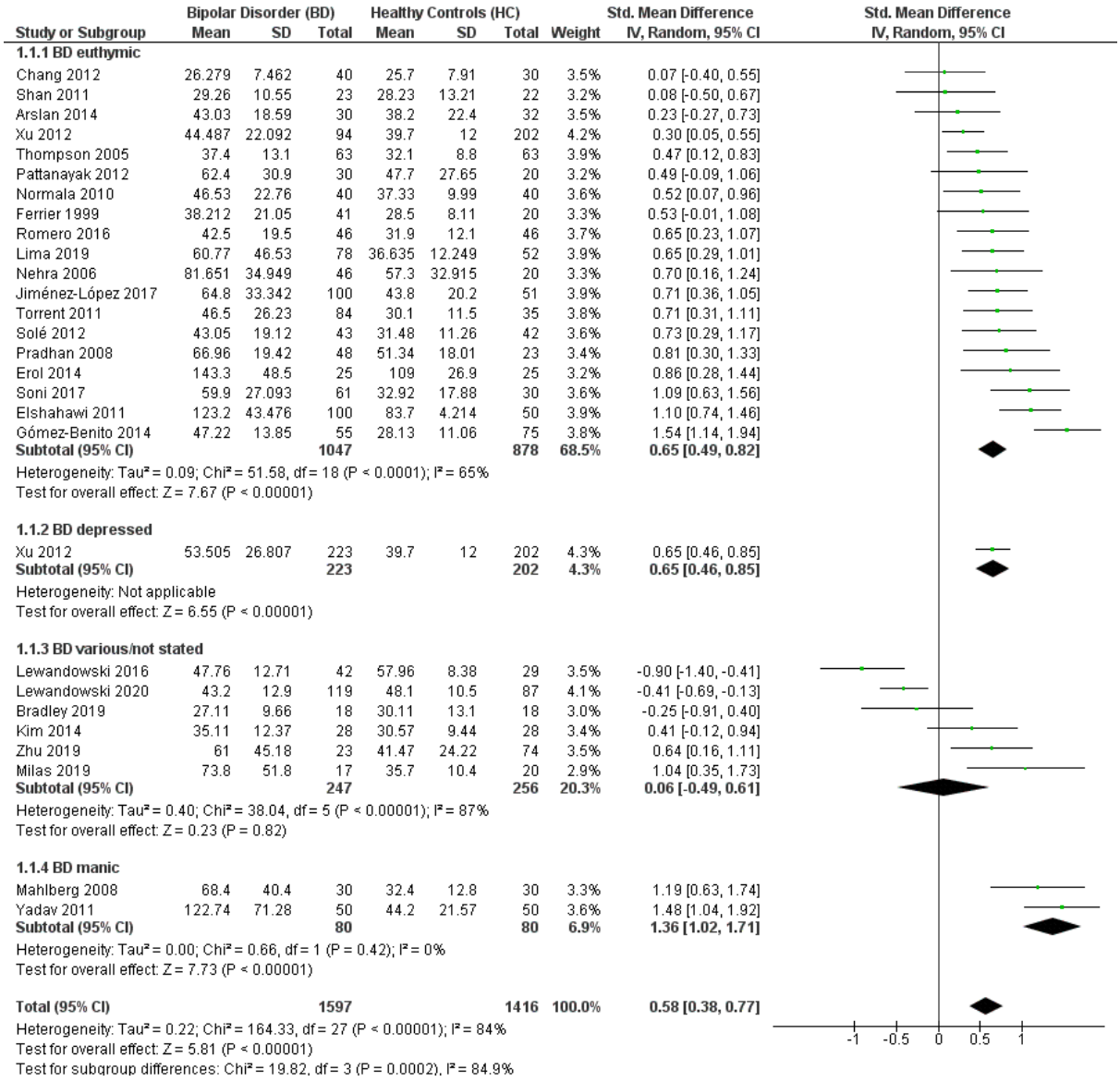

Figure 7: Forest plot of the comparison between BD patients and healthy controls on TMT-A time to complete.

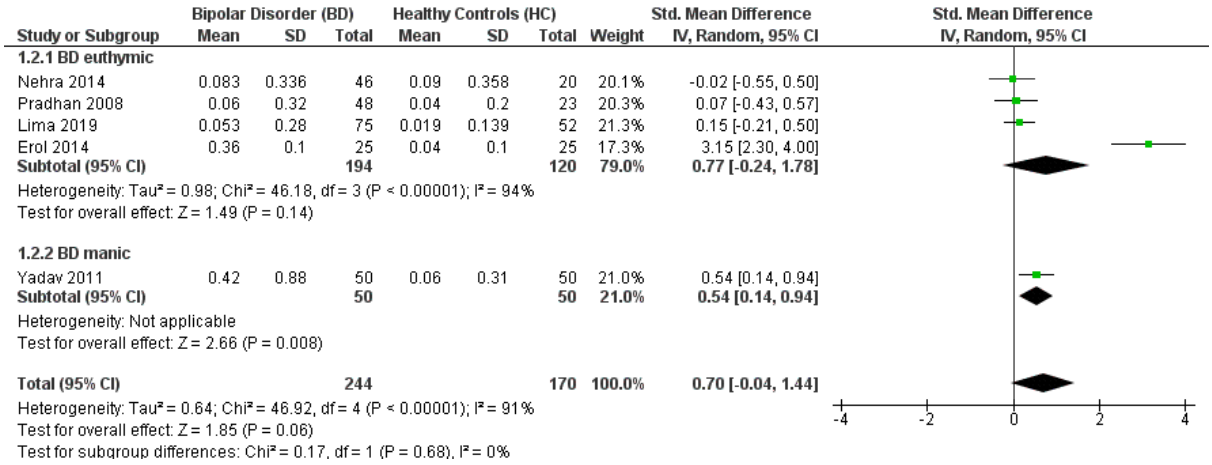

Figure 8: Forest plot of the comparison between BD patients and healthy controls on TMT-A number of errors.

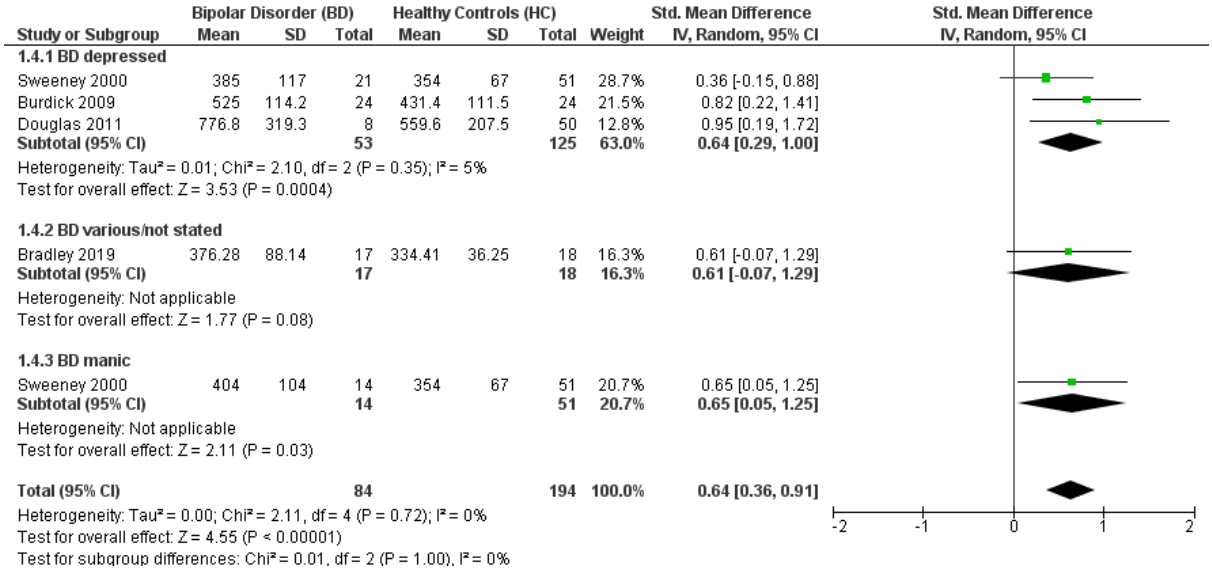

Figure 9: Forest plot of the comparison between BD patients and healthy controls on simple-RT tasks.

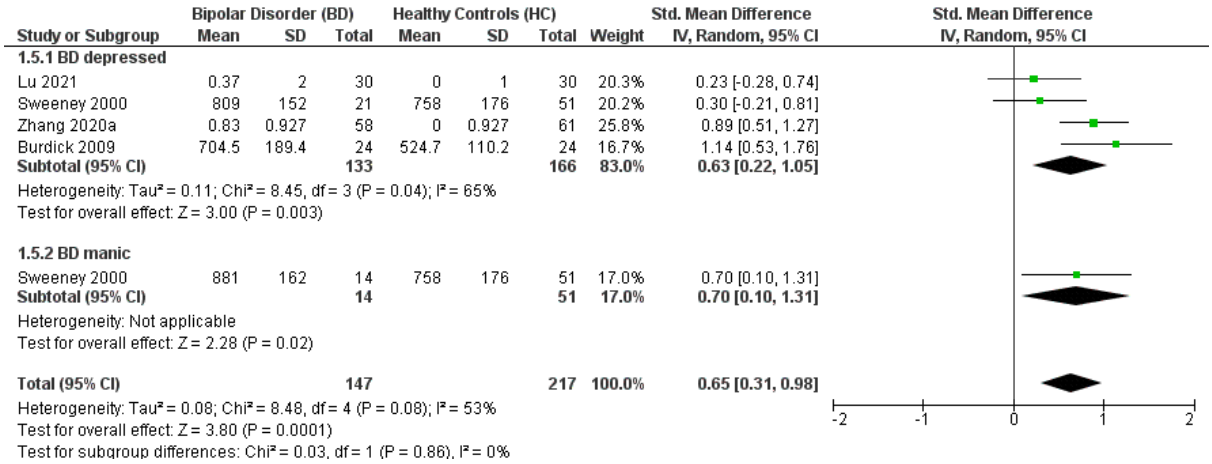

Figure 10: Forest plot of the comparison between BD patients and healthy controls on choice-RT tasks.

3.3.4.4. Verbal fluency

Verbal fluency tasks are generally split into two types: category fluency and phonemic fluency. We first analysed both tasks together, then split them by type to analyse category fluency tasks and phonemic fluency tasks separately. For studies that reported data for both types of fluency test, we combined the group means and SDs for the two types of test using the Cochrane formula (Higgins et al., 2022). Note that for this combined comparison, the number of correct responses was measured. Figure 11 shows the result of this comparison; BD patients provided fewer correct responses overall compared to controls (moderate ES). Eleven studies tested a sample of euthymic patients, *k*=2 tested a sample of depressed patients, and *k*=5 had an unspecified/mixed sample. The group difference was significant for each of these subgroups and there was no significant difference between the subgroups. Sensitivity analysis showed that removing any of these subgroups did not change the overall outcome. There was significant heterogeneity overall and for the euthymic and unspecified/mixed subgroups.

We also analysed the category fluency and phonemic fluency data separately. Two studies looked into number of errors on phonemic fluency tasks (Lima et al., 2019; Thompson et al., 2005), therefore this was not meta-analysed. Neither study showed a significant difference (ESs were $d<0.01$ and $d=0.16$, respectively). Meta-analysis was conducted for number of correct responses for both types of fluency tasks; the results of these comparisons are shown in Figure 12 and Figure 13. Note that some data reported here were disaggregated data from the combined comparison reported above. For category fluency, BD patients gave fewer correct responses overall compared to controls. Seven studies tested a sample of euthymic patients, and two studies had an unspecified/mixed sample. The overall effect was significant for the euthymic subgroup, but not the unspecified/mixed subgroup. However, the subgroup analysis showed that there was no significant difference between the subgroups. There was significant heterogeneity overall and for the unspecified/mixed subgroups, but the euthymic subgroup was not heterogeneous. For phonemic fluency, BD patients gave fewer correct responses overall compared to controls. Ten studies tested a sample of euthymic patients, *k*=2 had a sample of depressed patients, and *k*=1 had an unspecified/mixed sample. The overall effect was significant for each subgroup and there was no significant difference between the subgroups. There was significant heterogeneity overall and in the euthymic subgroup.

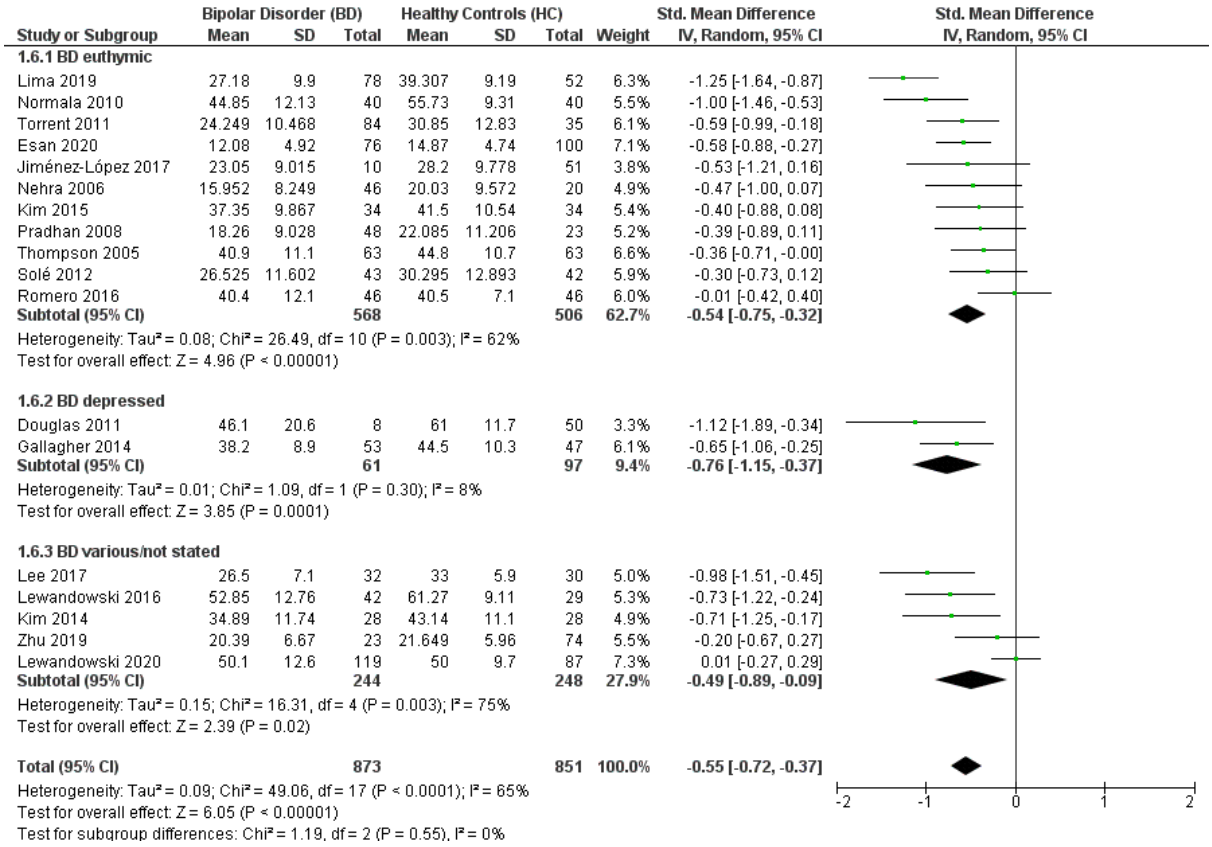

Figure 11: Forest plot of the comparison between BD patients and healthy controls on verbal fluency tasks (number of correct responses).

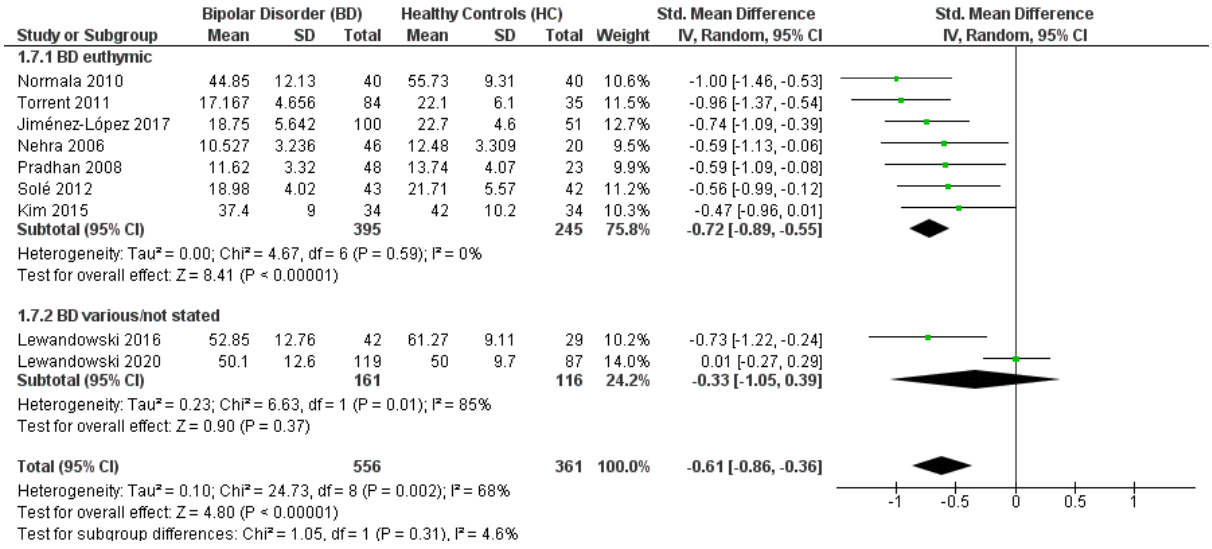

Figure 12: Forest plot of the comparison between BD patients and healthy controls on category fluency tests (number of correct responses).

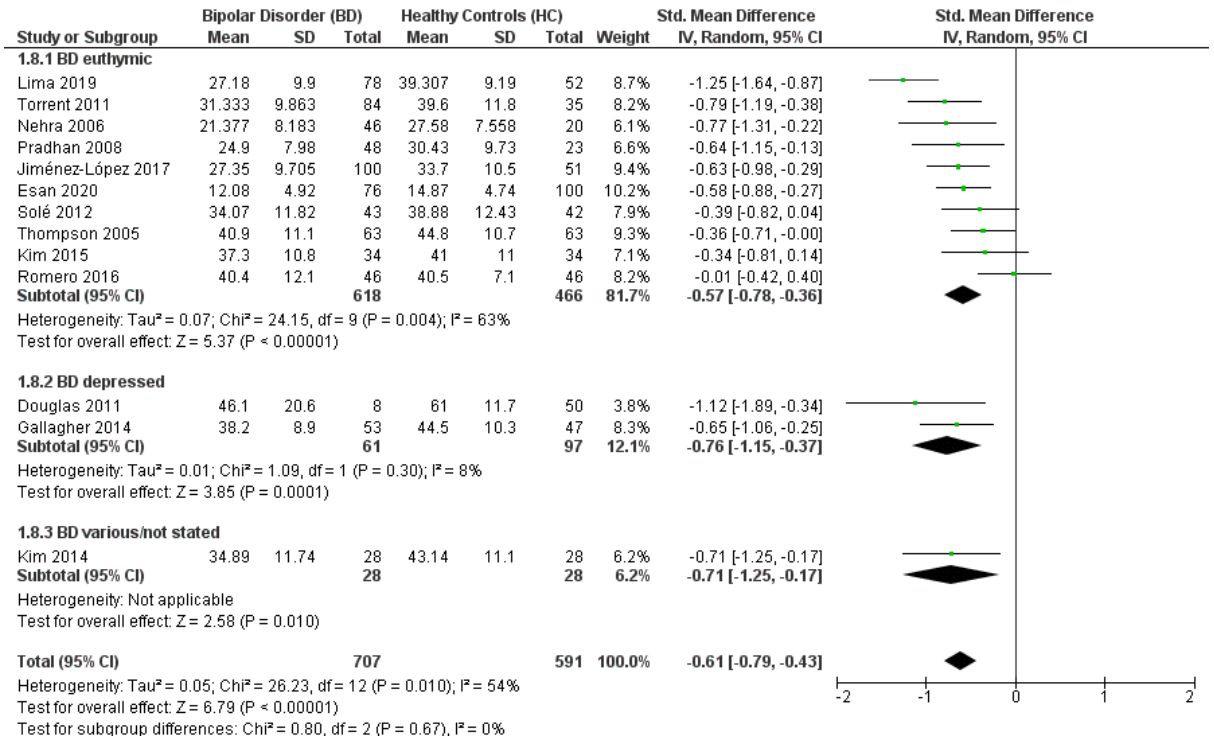

Figure 13: Forest plot of the comparison between BD patients and healthy controls on phonemic fluency.

3.3.4.5. Stroop task

The Stroop task usually consists of three parts or blocks: a block of colour-naming trials, a block of word naming trials, and a block of colour-word trials. The latter block is known as 'interference' trials, which require inhibition of the colour-naming response to be able to respond to the word and is considered a measure of EF. The former two blocks are simple/automatic trails and are comparable with simple-RT tasks. We therefore meta-analysed data from these types of trials as a measure of PS. In some cases, studies reported data from both types of automatic trials separately, so these scores were combined using the Cochrane formula. Studies tend to report either number of correct responses or time to complete as the outcome measure; we reversed time to complete scores so that higher scores reflect better performance, then pooled data using both outcome variables. Four studies had a sample of euthymic patients, one study had a sample of depressed patients, and one study had a sample with unspecified/mixed mood states. The results are shown in Figure 14; BD scored worse than controls overall (large ES, *d*=0.79) as well as in each subgroup. There was no difference between subgroups and there was no heterogeneity across the studies.

Since we combined two types of trials (colour naming and word reading), as well as two outcome variables (number correct and time to complete), we also meta-analysed these

variables separately. The forest plots of the results can be found in Appendix A (Figure 44, Figure 45, Figure 46, and Figure 47); that BD scored worse than HCs on all comparisons.

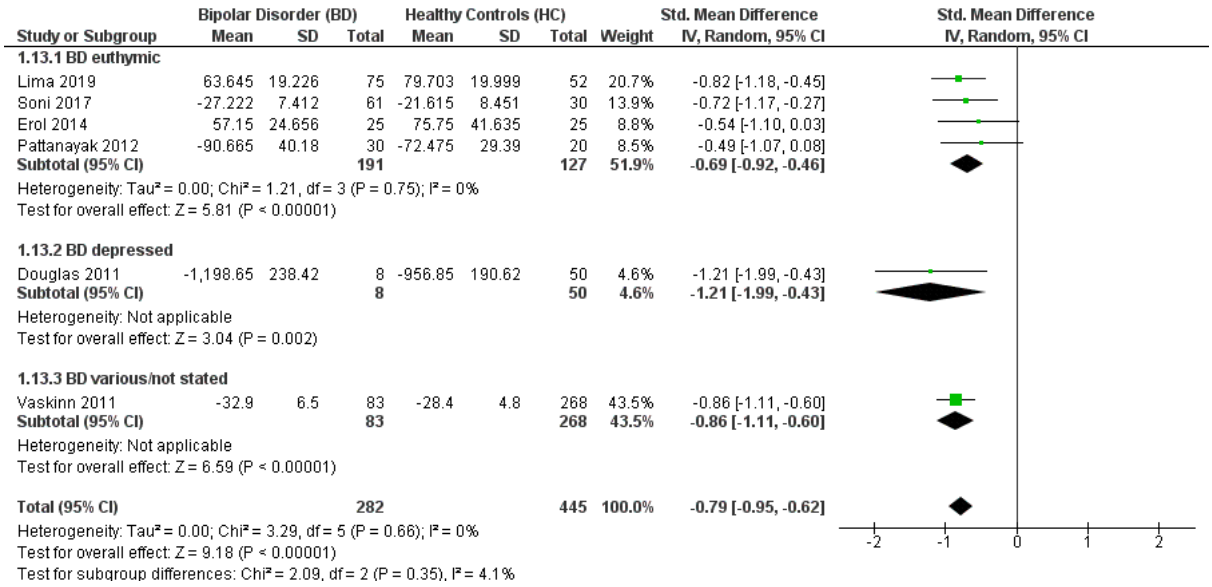

Figure 14: Forest plot of the comparison between BD patients and healthy controls on Stroop simple/automatic trials (word reading and colour naming combined; all outcome measures).

3.3.4.6. Other tests of PS

Four studies (*k*=3 with a euthymic sample, *k*=1 with an unspecified/mixed sample) reported composite scores for PS but did not provide disaggregated data for each neuropsychological outcome measure. We pooled the composite PS scores and the overall results showed that BD patients performed worse than controls (Figure 15). There was significant heterogeneity in the data, which appeared to be driven by the study with an unspecified/mixed sample.

Two studies reported scores from other tests of PS in BD depression: Douglas et al. (2011) used the Timed Chase Test as a measure of psychomotor speed. In this test, participants must chase a moving tile around a grid of squares on a computer screen for 30 seconds. There was no significant difference between BD and HC in the number of moves, however, there were only 8 BD participants in this study. Gallagher et al. (2014) used the Speed and Capacity of Language Processing (SCOLP) test to test the speed and efficiency of cognitive processing. The SCOLP has two scores, 'spot the word' and 'speed of processing', which are both measured by total correct responses. There was no significant difference between groups in the spot the word score, but BD performed worse than HCs on the speed of processing score.

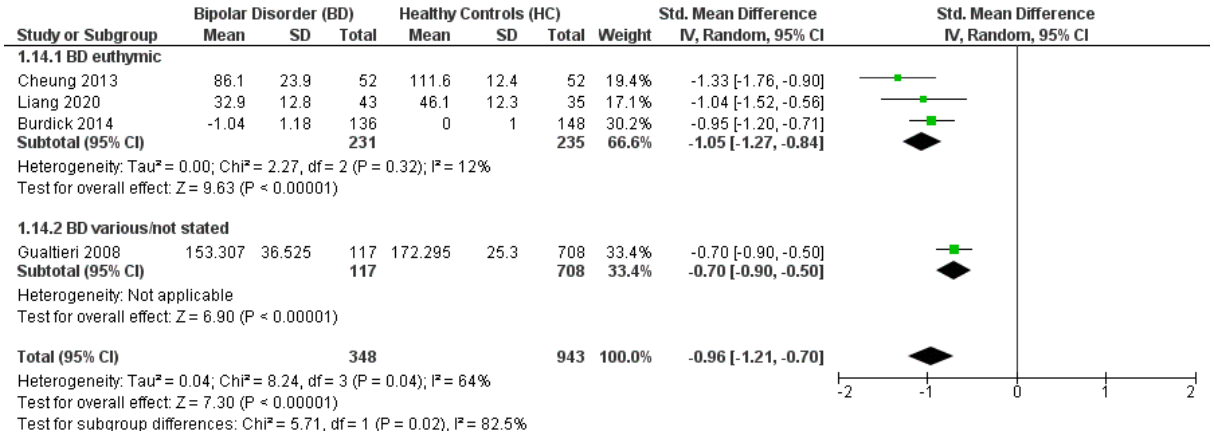

Figure 15: Forest plot of the comparison between BD patients and healthy controls on composite processing speed scores.

3.3.4.7. CPT accuracy/correct/errors

Eight studies reported accuracy/number correct scores on CPTs for BD patients. The overall result shows that BD participants perform less accurately than controls on CPTs (see Figure 16). Six studies included a sample of euthymic BD patients, and two studies did not specify the mood state of participants. There was no significant difference between subgroups, however the pooled effect for the unspecified sample was not significant (though this was only for two studies). There was no significant heterogeneity overall.

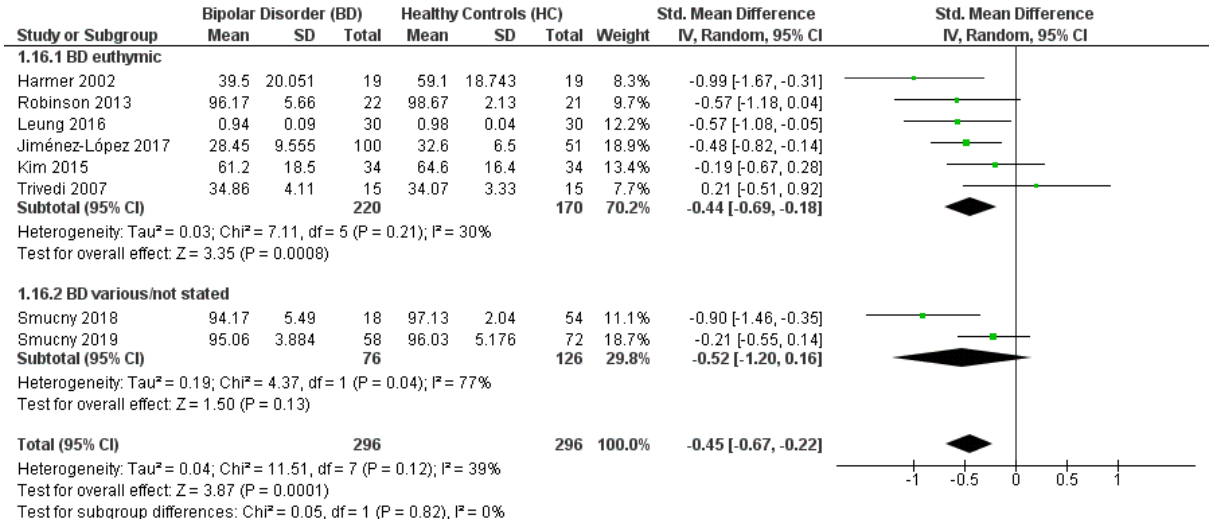

Figure 16: Forest plot of the comparison between BD patients and healthy controls on CPT number correct.

Eight studies reported the number of errors of omission on CPTs for BD patients. The overall result shows that BD groups committed more errors of omission than controls (see Figure 17). Five studies included a sample of euthymic BD patients, two studies included a sample of

depressed BD patients, and one study did not specify the mood state of participants. There was no significant difference between subgroups and there was no significant heterogeneity.

Nine studies reported the number of errors of commission on CPTs for BD patients. The overall result shows that BD groups committed more errors of commission than controls on CPTs (see Figure 18). Six studies included a sample of euthymic BD, and three studies included a sample of depressed BD. There was no significant difference between subgroups. There was significant heterogeneity overall and within each subgroup.

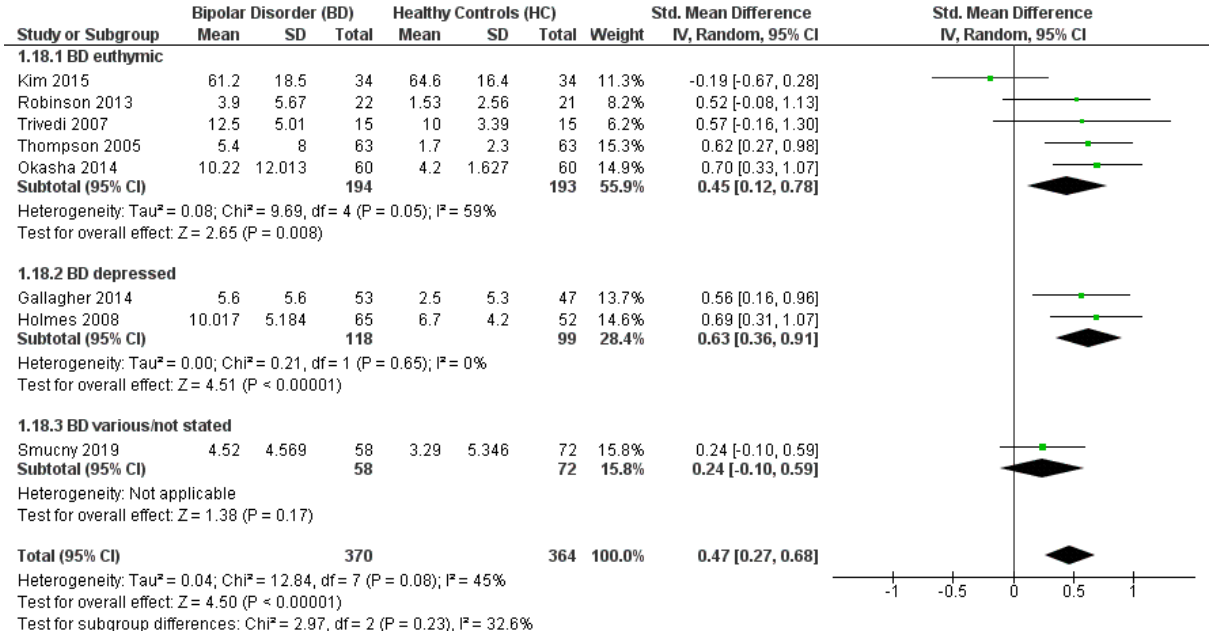

Figure 17: Forest plot of the comparison between BD patients and healthy controls on CPT errors of omission.

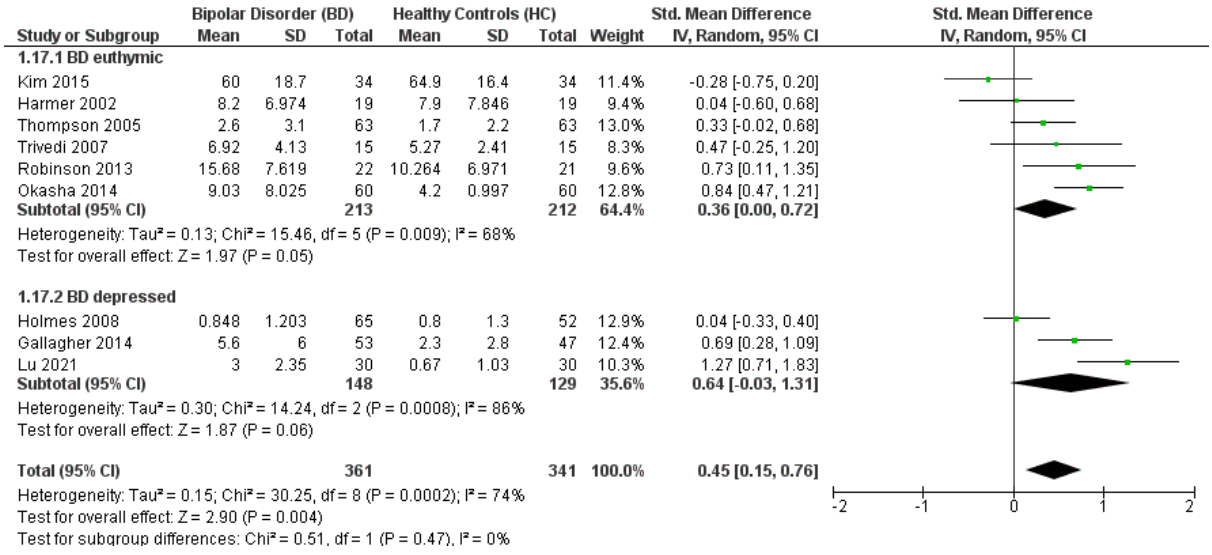

Figure 18: Forest plot of the comparison between BD patients and healthy controls on CPT errors of commission.

3.3.4.8. CPT sensitivity/d'

A sensitivity measure from signal detection theory (d') is often used to measure performance in CPTs. Twelve studies reported d' for BD patients: pooled data showed that BD patients performed worse than HCs (Figure 19). Five studies included a sample of euthymic BD, and seven studies did not specify the mood state of the sample. There was no difference between subgroups. There was significant heterogeneity overall and in the unspecified/mixed sample papers, which appeared to be driven one study (W. Kim et al., 2014); sensitivity analysis showed that when this was removed, there was no significant heterogeneity.

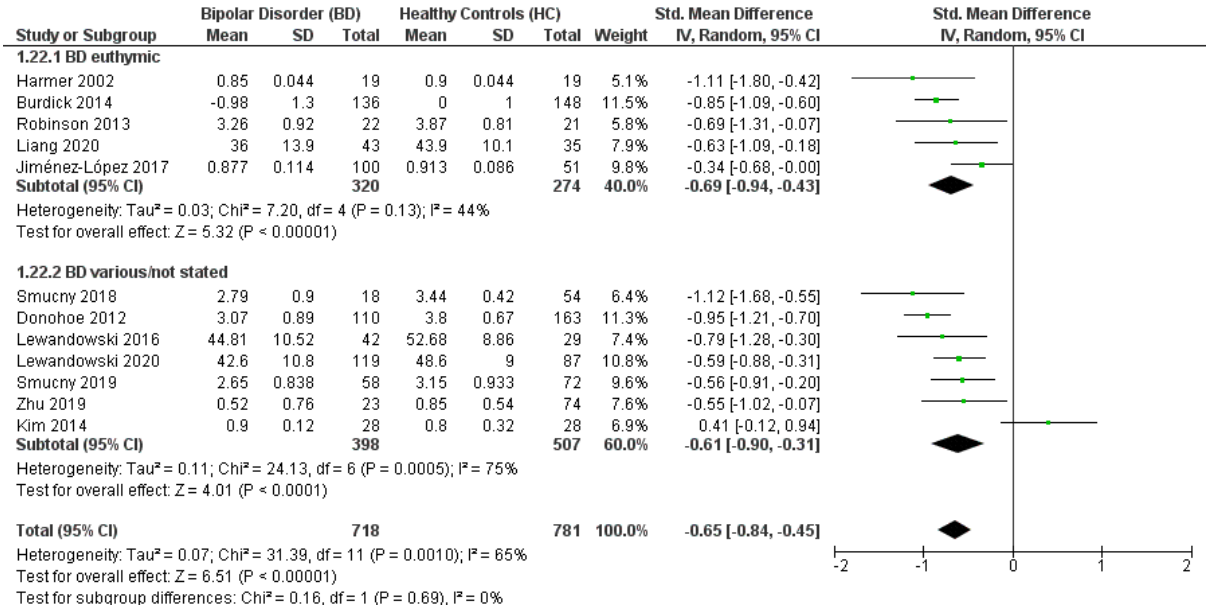

Figure 19: Forest plot of the comparison between BD patients and healthy controls on CPT sensitivity (d').

3.3.4.9. CPT average RT

Twelve studies reported CPT RTs for BD patients. The overall result was not significant (Figure 20). Seven studies included a sample of euthymic patients, *k*=3 included a sample of depressed patients, and *k*=2 did not specify the mood state of the sample. Subgroups were significantly different overall. Post-hoc pairwise subgroup tests suggested that the unspecified/mixed sample subgroup was significantly different from the euthymic sample (χ^2=6.11, *p*=.01), but no other pairwise comparisons showed a significant difference (both *p*s>.05). There was significant heterogeneity overall and within the euthymic and depressed subgroups. Two studies reported that euthymic patients had faster RTs than controls (Lima et al., 2019; Radwan et al., 2013). A sensitivity analysis showed that removing Lima et al. (2019) led to a significant overall result (*d*=.31, 95% CI=0.02 to 0.59, *Z*=2.13, *p*=.03).

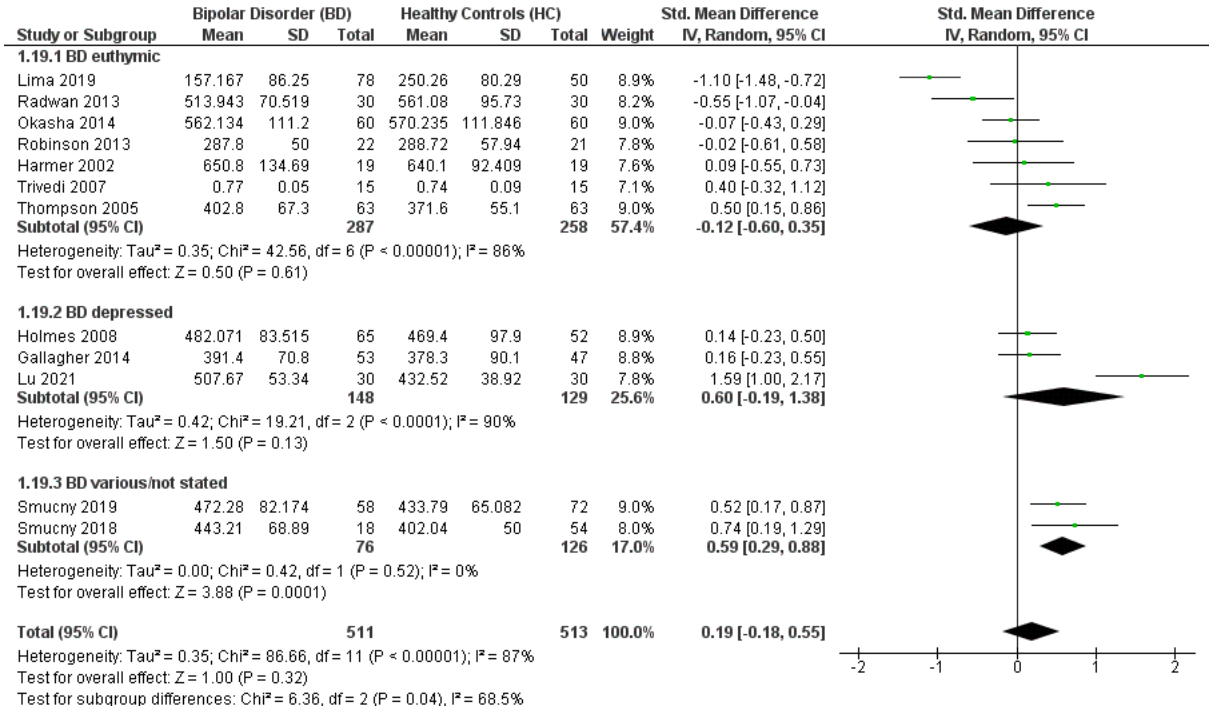

Figure 20: Forest plot of the comparison between BD patients and healthy controls on CPT average RT.

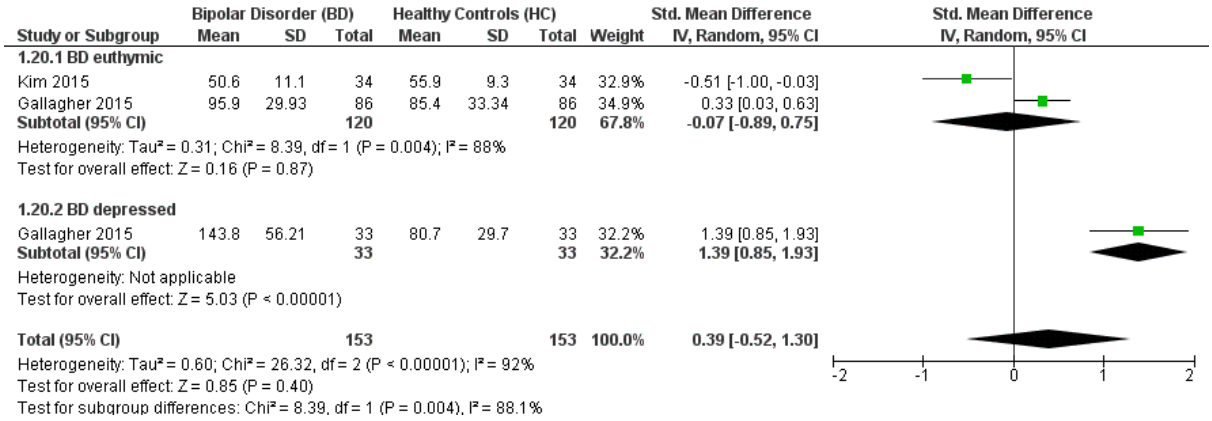

Figure 21: Forest plot of the comparison between BD patients and healthy controls on CPT RT variability (iSD).

3.3.4.10. CPT RT variability

Two studies reported metrics that assessed RT variability for CPT data (Gallagher, Nilsson, et al., 2015; D. Kim et al., 2015). Gallagher et al. reported iSD and CoV for a sample of depressed BD and a sample of euthymic BD. Kim et al. reported iSD data for a sample of euthymic BD. Data for iSD are shown in Figure 21. There was no overall difference between BD and HCs, however, only three samples from three studies were pooled here from two articles. The depressed and euthymic BD samples from Gallagher et al. appeared to have significantly larger iSD from controls, however the euthymic BD sample from Kim et al., appeared to have reduced

iSD compared to controls. For CoV, Gallagher et al. reported a significant increase in depressed BD compared to HCs, but there was no difference between euthymic BD and controls.

3.3.4.11. CPT other outcome variables

One study reported ex-Gaussian metrics for CPT RT for depressed and euthymic BD groups (Gallagher, Nilsson, et al., 2015). Neither euthymic or depressed BD samples showed a significant difference from controls in *mu*, only the depressed sample showed a significant difference from controls in *sigma*, and both euthymic and depressed patients showed a significant difference from controls in *tau*. Two studies reported response bias/criterion metrics from signal detection theory for CPT data for BD (Harmer et al., 2002; Robinson et al., 2013). Neither study found a significant difference between euthymic BD patients and controls.

3.3.4.12. Other tests of SA

Five studies used other measures of SA: Cheung et al. (2013) and Gualtieri and Johnson (2008) reported a composite score of attention from CNSVS, which was made up of number of errors in CPT, a shifting attention test, and the Stroop test. Cheung et al. included sample of euthymic BD; data showed that patients performed significantly worse than controls with a large ES (*d*=1.31). Gualtieri et al. included sample of BD with unspecified mood states; data showed that patients performed significantly worse than controls on this composite score with a large ES (*d*=0.83). Bradley et al. (2019) reported RTs from the Attention Network Test (ANT), and the data suggests that BD patients (mixed mood state) did not significantly differ from controls. Frydecka et al. (2014) reported data from a modified version of a CPT that included a WM component. The data suggested that there was no significant difference between BD patients (unspecified mood states) and controls. Marotta et al. (2015) reported scores from an ANTI-vigilance task, which is like ANT with an added vigilance component. They reported that performance, as measured by several vigilance indices, was lower in BD compared to controls (including accuracy, d', and RTs). However, percentage of false alarms and response bias were not significantly different.

3.3.5. Meta-analysis of major depressive disorder

3.3.5.1. DSST

The number of correct responses on the DSST (pooled across 90s and 2-minute test periods) was significantly lower for MDD patients than HCs (Figure 22). Samples were grouped based on mood state, including depressed (*k*=18), remitted (*k*=4), and samples from studies that did not specify the mood state of participants or contained mixed samples (*k*=6). The overall group difference was significant in the depressed and unspecified subgroups (*p*s <.05), but not for the remitted groups (*p*=.06). There was no significant difference between subgroups overall (*p*=.06), but when depressed and remitted groups were compared, there was a significant group difference (*p*=.02). There was significant heterogeneity overall, as well as within each subgroup. Only one study reported another metric from the DSST: MDD patients tended to make more errors on DSST than HCs, but this was not statistically significant (Pier et al., 2004).

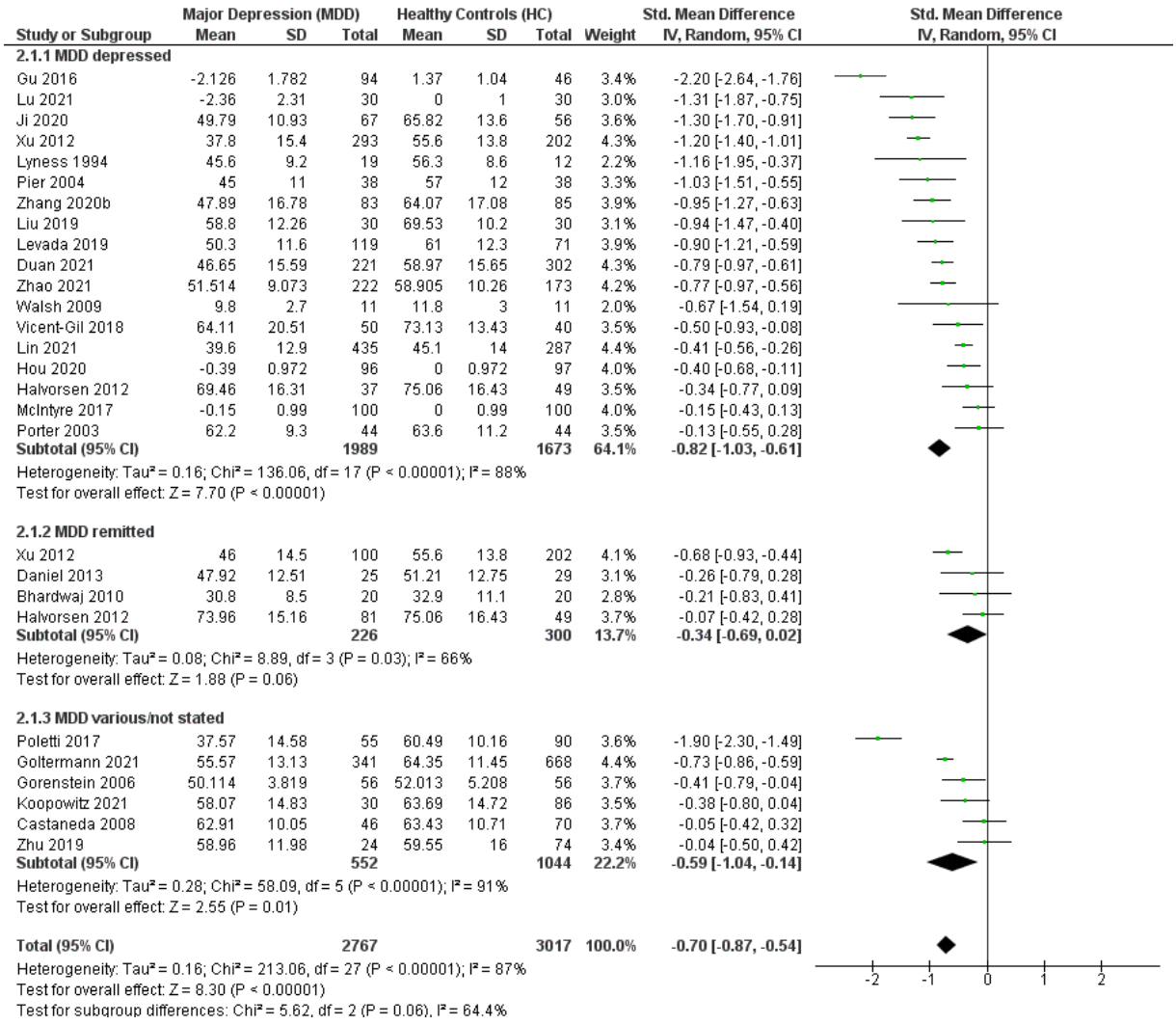

Figure 22: Forest plot of the comparison between MDD patients and healthy controls on DSST (number correct).

3.3.5.2. TMT-A

Most studies that reported TMT-A scores measured the time to complete the task as the outcome variable. Time taken to complete the TMT-A was significantly higher for MDD patients than controls, as shown in Figure 23. For this comparison, there were enough studies to group samples based on mood state, including depressed (*k*=12), remitted (*k*=5), and unspecified/mixed samples (*k*=5). The overall group difference was significant in each of the subgroups (all *p*s <.05). Subgroups significantly differed overall (*p*=.001). Post-hoc pairwise subgroup tests showed that the depressed and not specified/mixed subgroups significantly differed (*p*<.001), but the other two comparisons were not significant (both *p*s>.05). There was significant heterogeneity overall as well as within the depressed and remitted subgroups.

Only two studies used number of errors as an outcome measure for TMT-A; one included a sample of remitted patients (Yamamoto & Shimada, 2012), and one included a sample of depressed patients (Halappa et al., 2018). Neither study reported a significant difference in number of errors of MDD patients compared to HCs. Two studies reported other outcome variables from the TMT-A: Halvorsen et al. (2012) reported number of correct responses and found no difference between depressed patients and controls; and Bakusic et al. (2021) used percentile scores and found that depressed patients scored worse than controls (*d*=0.74).

3.3.5.3. Reaction time

Performance on simple-RT tasks was not significantly worse for MDD patients than controls, see Figure 24. Four studies reported data for depressed samples and one study reported data for remitted MDD patients. A sensitivity analysis showed that removing the remitted study led to a significant overall result (*d*=0.40, 95% CI=0.04 to 0.77, *Z*=2.18, *p*=.03). There was significant heterogeneity overall and in the depressed subgroup.

Overall, MDD had longer RTs on choice-RT tasks than HCs (*p*<.001), see Figure 25. Six studies included a sample of depressed patients, two included a sample of remitted patients, and two did not specify whether patients were depressed or remitted. There was no significant subgroup difference and no pair-wise subgroup differences (all *p*s>.05). However, only the depressed subgroup had a significant overall effect. There was significant heterogeneity overall and in the depressed subgroup.

3.3.5.4. Verbal fluency

We first analysed category fluency and phonemic fluency tasks together (Figure 26). MDD patients gave fewer correct responses overall compared to controls. Ten studies tested a depressed sample, *k*=3 had a remitted sample, and *k*=1 had an unspecified/mixed sample. The group difference was significant for the depressed and remitted subgroups, and there was no significant difference between the subgroups. Sensitivity analysis showed that removing any of these subgroups did not change the overall outcome. There was significant heterogeneity for the depressed subgroup; the overall heterogeneity was at *p*=.05.

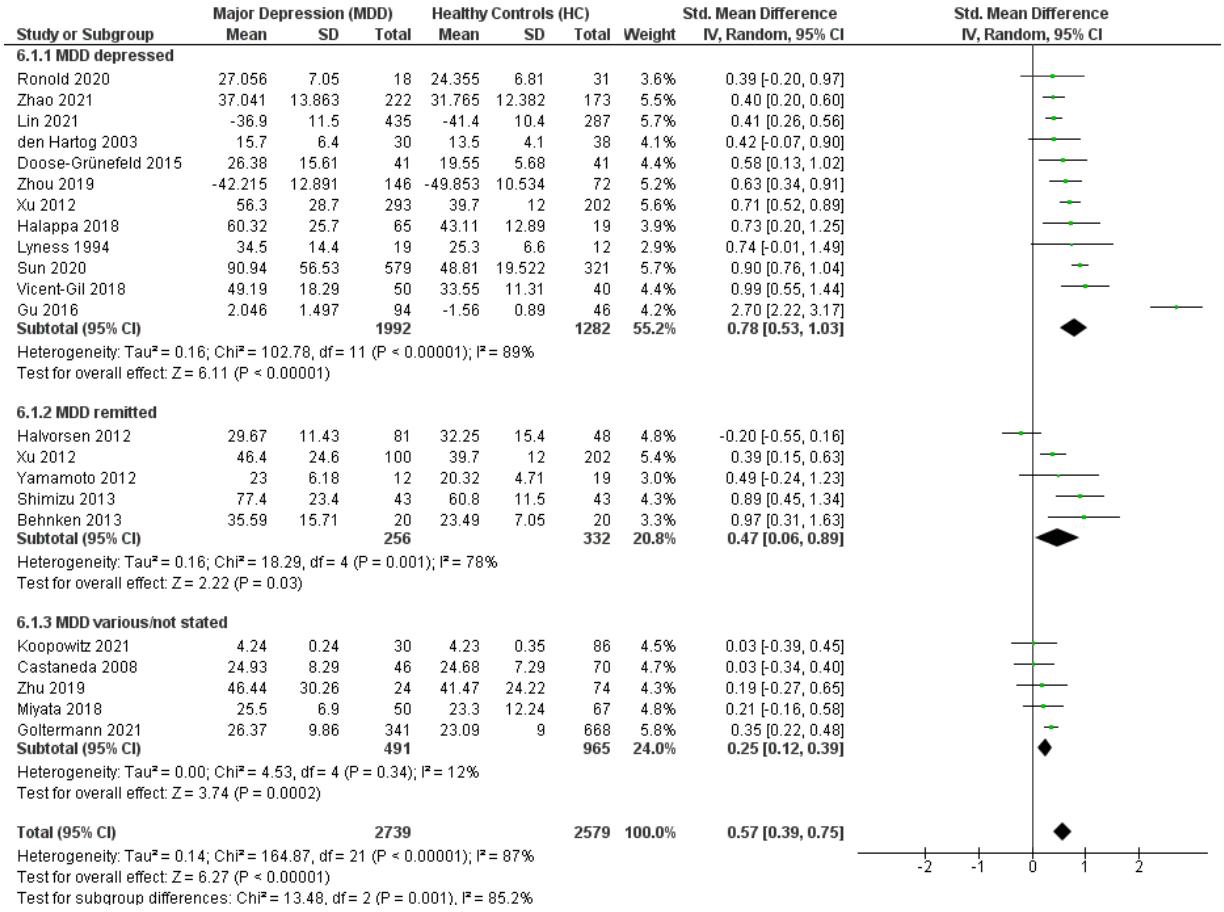

Figure 23: Forest plot of the comparison between MDD patients and healthy controls on TMT-A time to complete.

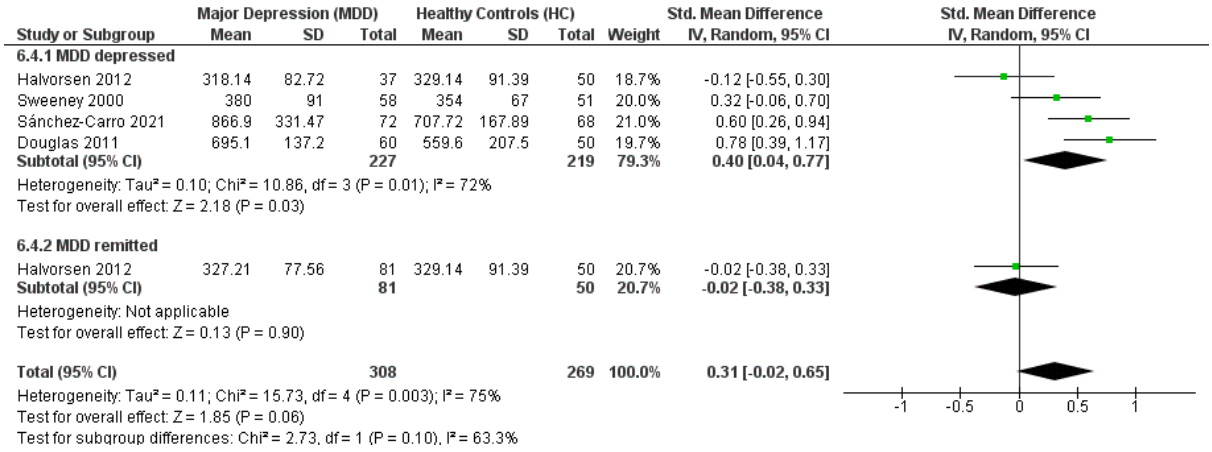

Figure 24: Forest plot of the comparison between MDD patients and healthy controls on simple-RT tasks.

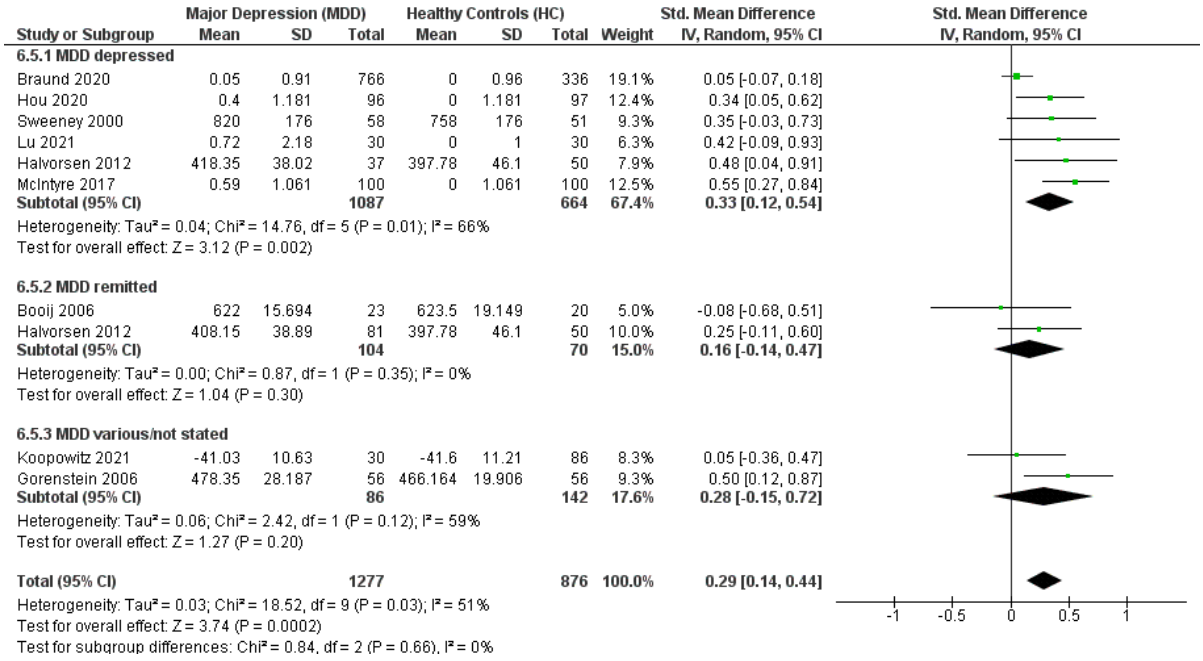

Figure 25: Forest plot of the comparison between MDD patients and healthy controls on choice-RT tasks.

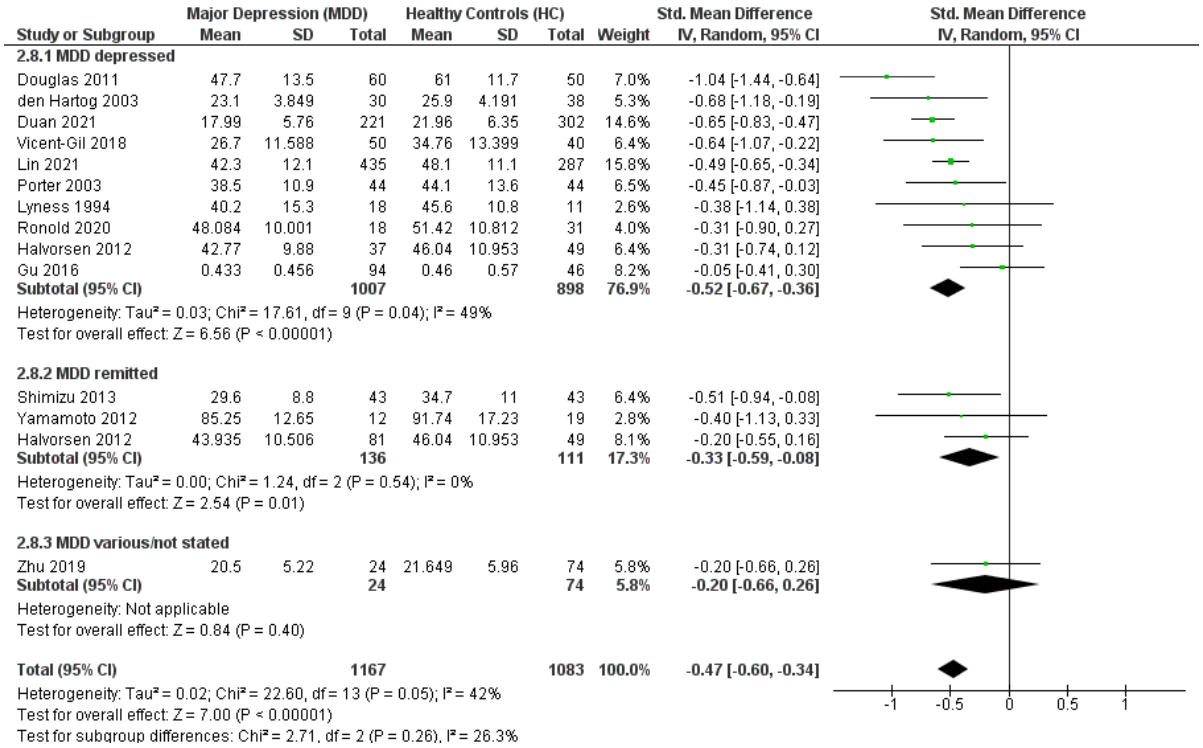

Figure 26: Forest plot of the comparison between MDD patients and healthy controls on verbal fluency tasks.

We also analysed the category fluency and phonemic fluency data separately (Figure 27 and Figure 28, respectively). Note that some data reported here were disaggregated data from the combined comparison. Overall, MDD gave fewer correct category fluency responses than HCs. *k*=7 tested a depressed sample, *k*=1 had a remitted sample, and *k*=1 did not specify the mood state of participants. The overall effect was significant for the depressed subgroup, but not for the remitted or unspecified/mixed subgroups. There was significant heterogeneity overall and for the depressed subgroup. However, subgroup analysis showed no significant differences between the subgroups. MDD gave fewer correct phonemic fluency responses compared to controls. *k*=6 included a depressed sample and *k*=2 included a remitted sample. The overall difference was significant for the depressed subgroup, but not the remitted subgroup, however, there was no significant difference between the subgroups. There was significant heterogeneity overall and in the depressed subgroup.

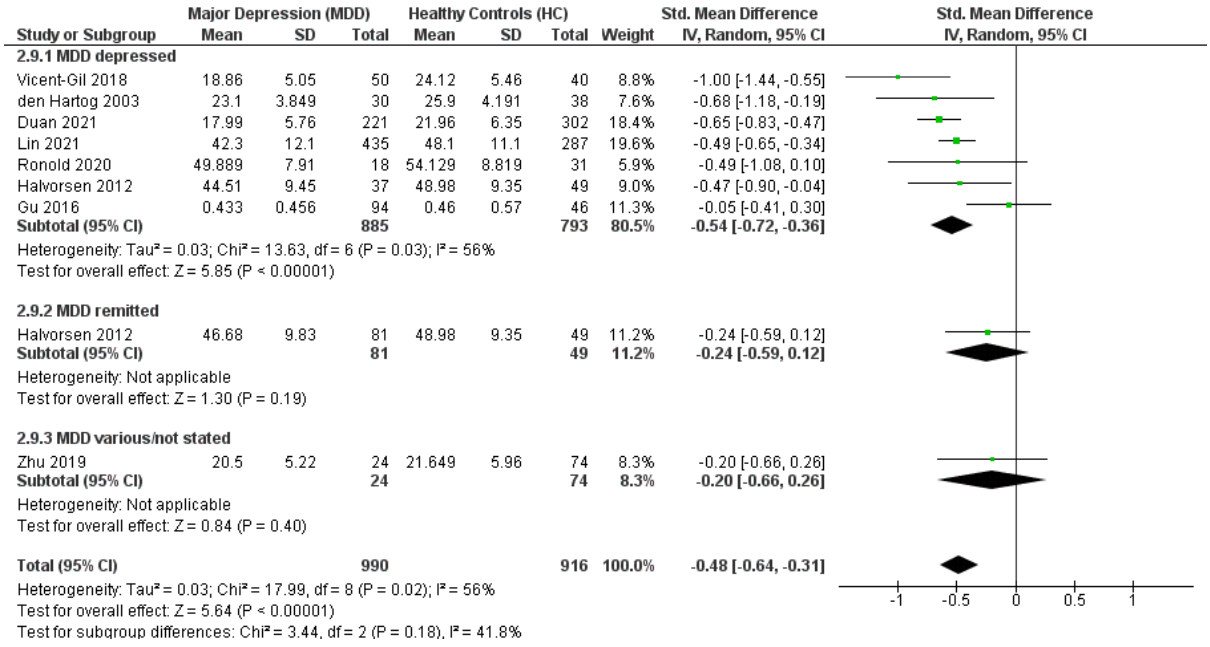

Figure 27: Forest plot of the comparison between MDD patients and healthy controls on category fluency tasks.

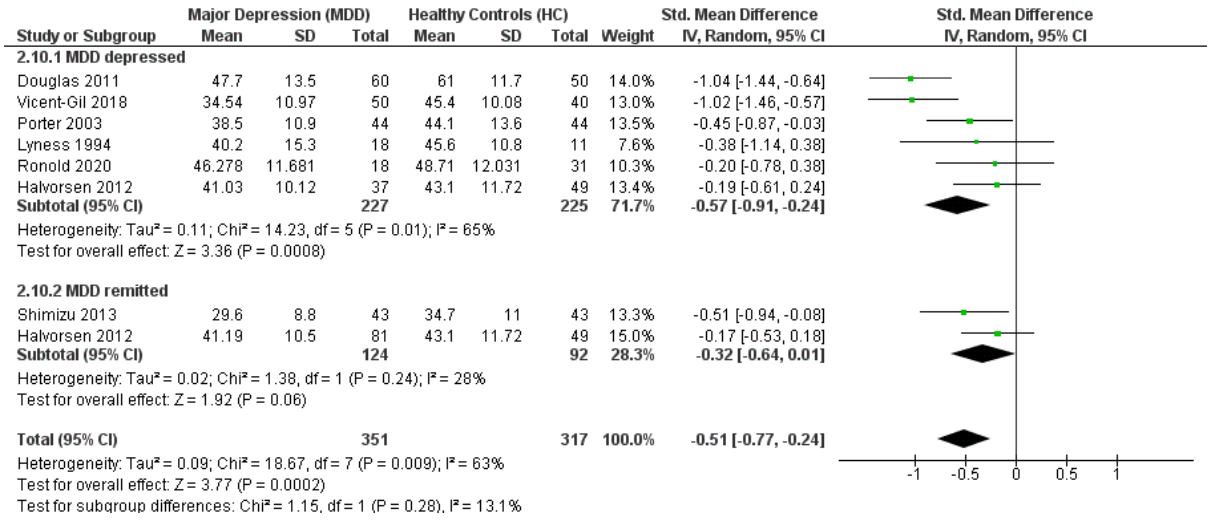

Figure 28: Forest plot of the comparison between MDD patients and healthy controls on phonemic fluency.

3.3.5.5. Stroop task

Figure 29 shows results for Stroop automatic/simple trials (colour naming and word reading combined). MDD scored worse than controls overall and in the depressed subgroup. The two studies with a remitted sample did not show a significant difference compared to controls. There was a significant difference between subgroups and significant heterogeneity overall and in the depressed subgroup. We also meta-analysed the colour naming and word reading trials separately; the forest plots detailing the results can be found in Appendix A (Figure 48 and Figure 49). The results showed the same pattern as the combined scores: MDD scored

worse than controls on both colour naming and word reading overall and in the depressed subgroup, but a significant difference was not found for remitted samples.

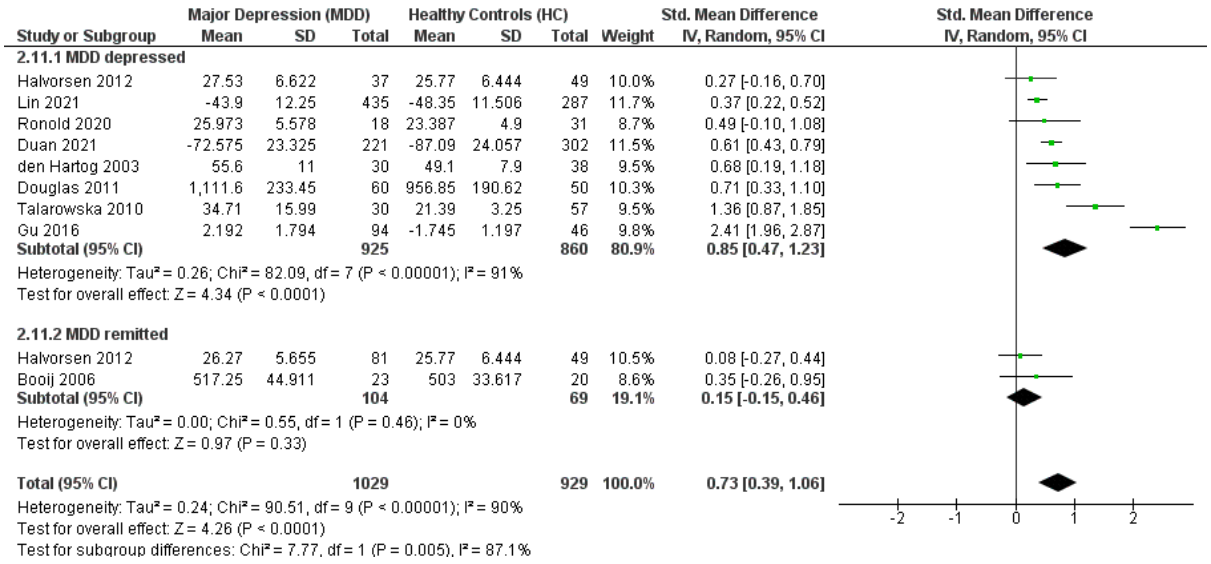

Figure 29: Forest plot of the comparison between MDD patients and healthy controls on Stroop simple trials.

3.3.5.6. Other tests of PS

Five studies reported composite scores of PS: (Jin et al., 2020; Liang et al., 2020; Zhou et al., 2019) all studies tested depressed patients and used the MCCB PS composite score, which is comprised of TMT-A, BACS symbol coding, and category fluency scores. Gualtieri and Johnson (2008) did not state the current mood of the sample and used the CNSVS battery psychomotor speed composite score, which is comprised of finger tapping test, and total correct on Symbol Digit Copy. Fernández-Sevillano et al. (2021) tested a depressed MDD sample and used a composite score comprising Weschler Adult Intelligence Scale (WAIS-IV) Symbol Search and Symbol Coding. The meta-analysis of these studies is presented in Figure 30 and shows that MDD perform worse than HCs on PS composite scores, with a large ES (*d*=0.86). The result was significant in the depressed subgroup, as well as for the paper with an unspecified group. There was significant heterogeneity and the unspecified sample appeared to be significantly different from the depressed MDD studies. One study reported scores from another test of PS in patients with depressed MDD: Douglas et al. (2011) found that MDD performed worse than controls at the Timed Chase Test (*d*=-0.85, 95% CI=-1.24 to -0.45, *p*<.001).

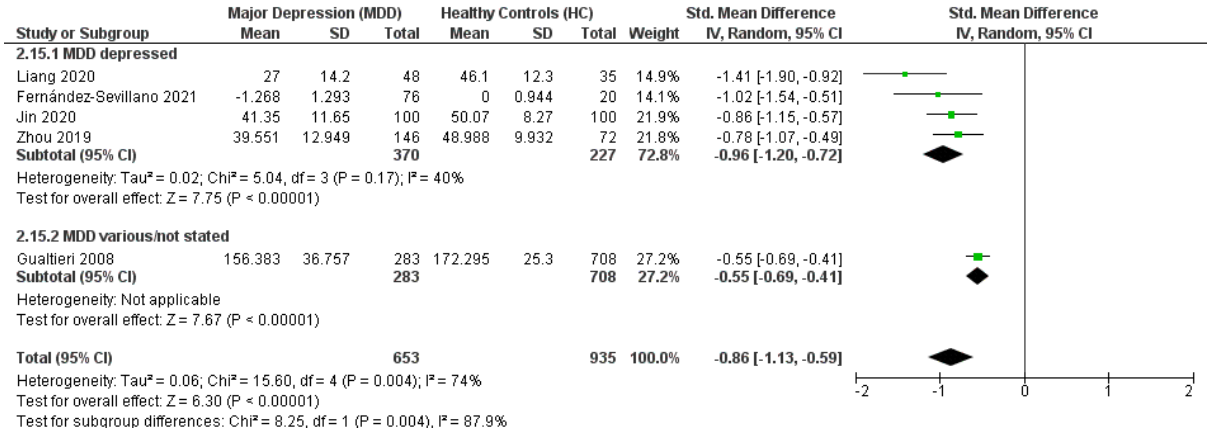

Figure 30: Forest plot of the comparison between MDD patients and healthy controls on PS composite scores.

3.3.5.7. CPT accuracy/number of errors

Only one study reported the number of correct responses for a CPT for MDD patients, and found that MDD had significantly fewer correct responses than controls (*d*=0.43, 95% CI=0.09 to 0.76, *p*=.01) (Sánchez-Carro et al., 2021). Four studies with a depressed sample reported the number of errors of omission of a CPT and *k*=5 with a depressed sample reported the number of errors of commission. Overall, MDD committed more errors of omission and commission than controls and there was no heterogeneity in either analysis (Figure 31 and Figure 32). One study reported total errors from the CPT for remitted MDD patients and found that there was no difference between remitted MDD and controls (*d*=0.20, 95% CI=-0.22 to 0.63, *p*=.35) (Shimizu et al., 2013).

3.3.5.8. CPT sensitivity/d'

Seven studies reported d' for CPTs; *k*=5 had a sample of depressed patients, and *k*=2 did not specify the mood state of participants (Figure 33). The overall result was significant, showing that MDD performed worse than controls on this measure. The result was significant for the depressed subgroup but not for the unspecified mood state subgroup. There was no significant heterogeneity overall, but there was significant heterogeneity in the depressed subgroup.

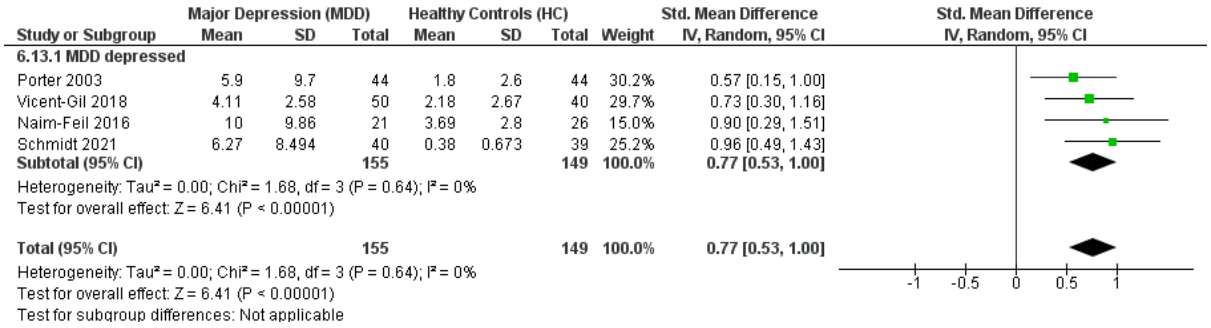

Figure 31: Forest plot of the comparison between MDD patients and healthy controls on CPT errors of omission.

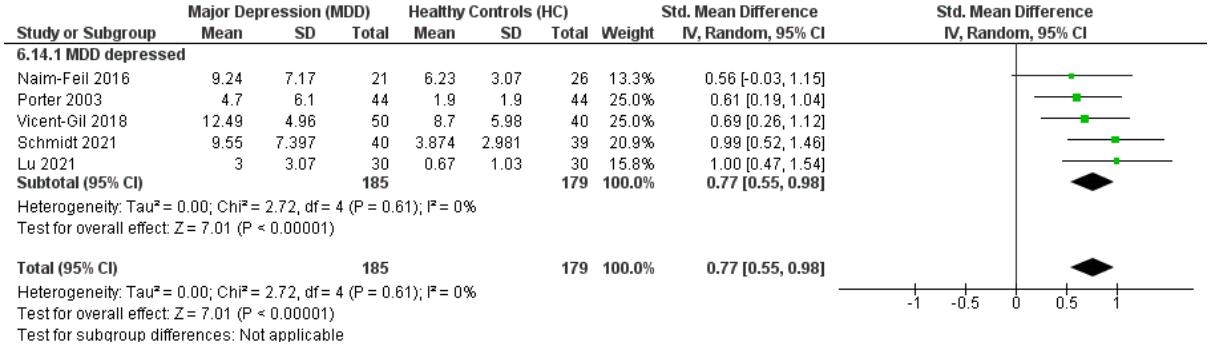

Figure 32: Forest plot of the comparison between MDD patients and healthy controls on CPT errors of commission.

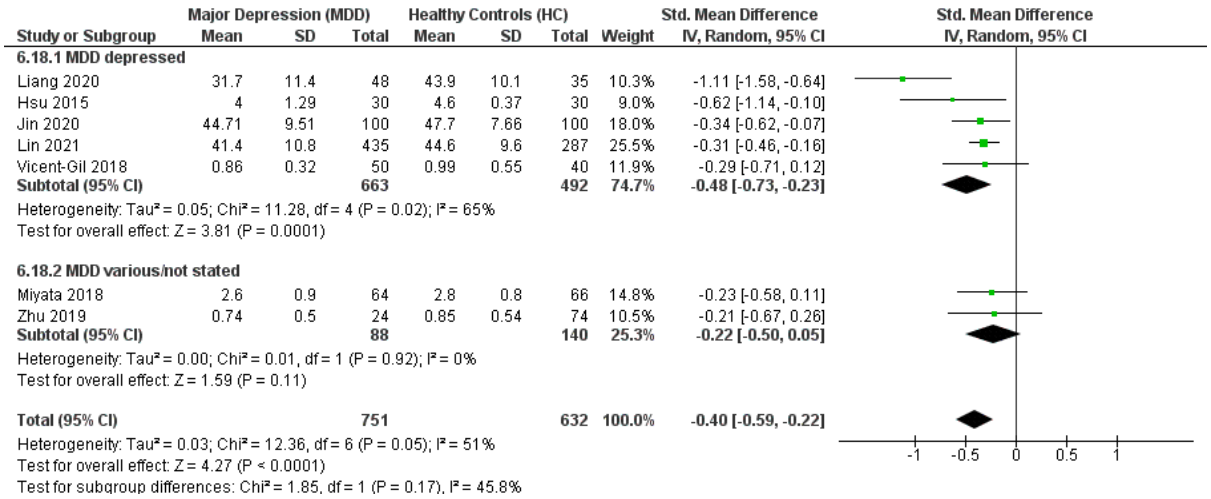

Figure 33: Forest plot of the comparison between MDD patients and healthy controls on CPT d'.

3.3.5.9. CPT average RT

For CPT RTs, MDD tended to have longer RTs than controls overall, however this was not significant (*p*=.05). There was high heterogeneity across these studies, with only three of the six studies showing significantly slower RTs than controls. *k*=5 included a sample of depressed MDD patients, and one study included a sample of remitted patients. A sensitivity analysis

suggested that removing the sample from Vicent-Gil et al. (2018) led to a significant overall result (*d*=0.80, 95% CI=0.27 to 1.33, *Z*=2.97, *p*=.003).

3.3.5.10. CPT RT variability

Few studies investigated variability of CPT RTs in MDD. Three studies reported iSD for depressed MDD: the overall result showed that MDD patients have more variable RTs than HCs (Figure 35). There was no significant heterogeneity across studies. *k*=3 reported CoV for MDD; two contained a sample of depressed MDD and one contained a sample of remitted MDD patients. The ES appeared moderate (*d*=0.67), however, the overall result not significant (*p*=.05) and there was significant heterogeneity across the studies (Figure 36).

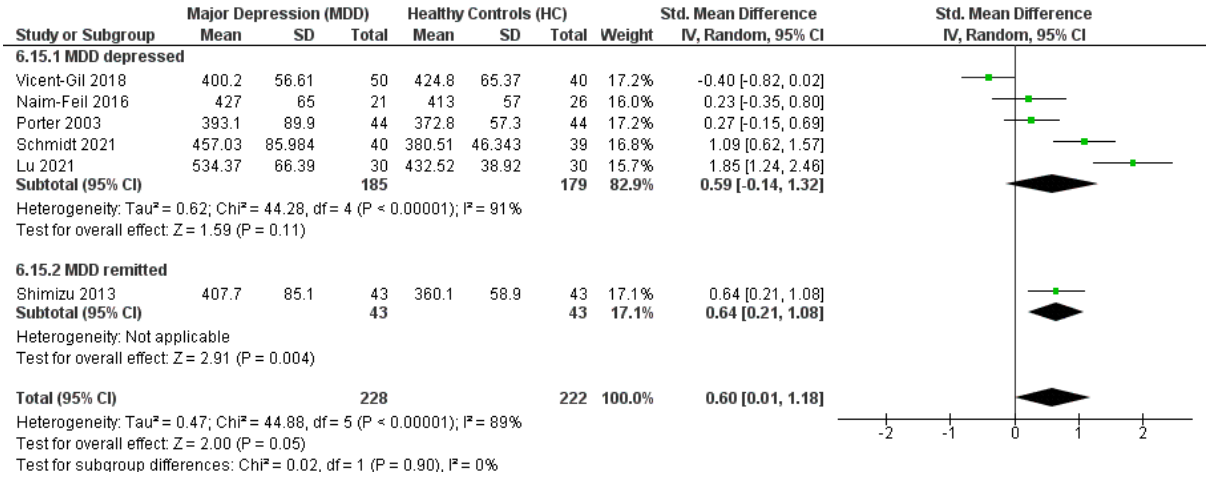

Figure 34: Forest plot of the comparison between MDD patients and healthy controls on CPT average RT.

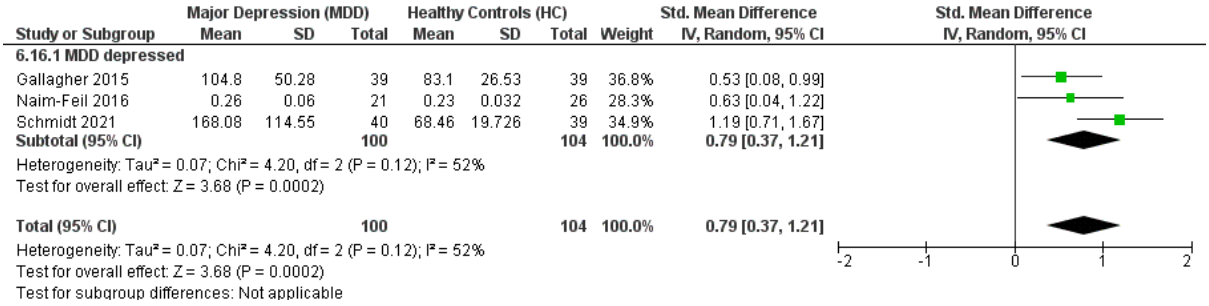

Figure 35: Forest plot of the comparison between MDD patients and healthy controls on CPT iSD (of RT).

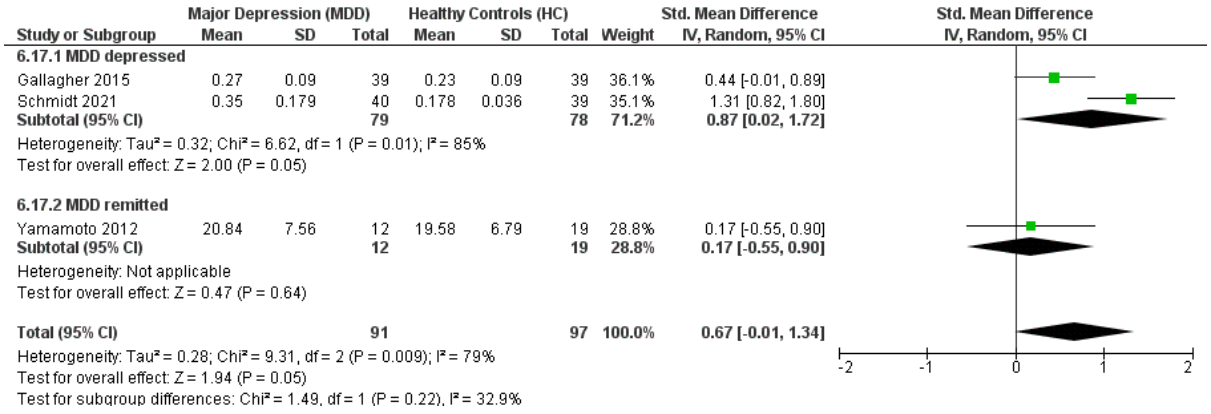

Figure 36: Forest plot of the comparison between MDD patients and healthy controls on CPT coefficient of variation (of RT).

3.3.5.11. CPT other outcome variables

Braund *et al.* (2020) reported a composite score using several CPT outcome measures, including total correct, false positive errors, false negative errors, RT, and RT variability. They found that depressed MDD patients had a significantly worse score than controls (*d*=-0.53, 95% CI= -0.66 to -0.40, *p*<.001). Only one paper utilised ex-Gaussian metrics to analyse CPT performance in depressed MDD patients (Gallagher, Nilsson, et al., 2015): they found no significant difference in *mu*, *sigma*, or *tau* in MDD patients compared to healthy controls (all *p*s>.05).

3.3.5.12. Other tests of SA

Two studies reported other tests of SA in MDD patients. Gualtieri and Johnson (2008) reported the CNSVS battery complex attention composite score in MDD (current mood not stated) and found that patients scored worse than controls (*d*=-0.56, 95% CI= -0.70 to -0.42, *p*<.001). Yamamoto and Shimada (2012) reported PASAT scores in remitted MDD and found that patients scored worse than controls (*d*=-1.11, 95% CI=-1.90 to -0.33, *p*=.005).

3.3.6. Heterogeneity

The heterogeneity statistics are summarised for each comparison in Table 3. Not all comparisons were heterogeneous: simple-RT and choice-RT, Stroop tasks, and CPT number correct and errors of omissions for BD; and CPT errors of omission and commission, sensitivity and iSD for MDD were not significantly heterogeneous. High heterogeneity was found for DSST, TMT-A, CPT RT, and CPT variability for both BD and MDD. High heterogeneity was found for other SA tasks for BD, and for Stroop tasks for MDD. The remaining comparisons had moderate heterogeneity.

	Bipolar disorder				Major depressive disorder			
Neuropsychological Outcome Variable	***Tau²***	***Q***	***p***	***I²***	***Tau²***	***Q***	***p***	***I²***
DSST (correct)	0.22	196.68	<.001	85.8	0.16	213.06	<.001	87.3
TMT-A (time)	0.22	164.33	<.001	83.6	0.14	164.87	<.001	87.3
TMT-A (errors)	0.64	46.92	<.001	91.5				
Simple-RT	0.00	2.11	.715	0.0	0.11	15.73	.003	74.6
Choice-RT	0.08	8.48	.076	52.8	0.03	18.52	.030	51.4
Verbal fluency (correct)	0.09	49.06	<.001	65.3	0.02	22.60	.047	42.5
Category fluency	0.10	24.73	.002	67.7	0.03	18.00	.021	55.5
Phonemic fluency	0.06	26.23	.010	54.3	0.09	18.7	.009	62.5
Stroop simple trials	0.00	3.29	.655	0.0	0.24	90.51	<.001	90.1
Stroop colour naming	0.00	1.83	.872	0.0	0.09	28.81	<.001	79.2
Stroop word reading	0.00	3.69	.450	0.0	0.07	26.35	<.001	73.4
PS composite	0.04	8.24	.041	63.6	0.06	15.60	<.001	74.4
CPT correct	0.04	11.51	.118	39.2				
CPT errors of commission	0.15	30.25	<.001	73.6	0.00	2.72	.605	0.0
CPT errors of omission	0.04	12.84	.076	45.5	0.00	1.68	.641	0.0
CPT sensitivity (d')	0.07	31.40	.001	65.0	0.03	12.36	.054	51.4
CPT RT	0.35	86.66	<.001	87.3	0.47	44.88	<.001	88.9
CPT RT variability (iSD)	0.60	26.32	<.001	92.4	0.07	4.20	.123	52.4
CPT RT variability (CoV)					0.28	9.31	.009	78.5
Sustained attention other	0.19	15.99	.001	81.2				

Table 3: Heterogeneity of each meta-analysis comparison. DSST=Digit Symbol Substitution Test; TMT=Trail-Making Test; RT=reaction time; PS=processing speed; CPT=continuous performance test; iSD=individual standard deviation; CoV=coefficient of variation.

3.3.7. Sensitivity analyses

A sensitivity analysis was performed for each comparison, where each study was removed individually to see if it changed the overall result of each comparison. For most comparisons, the statistical significance of the overall result remained unchanged no matter which study was removed from the analysis. For BD, removing Lima et al. (2019) from the CPT RT analysis changed the result from non-significant to significant. For MDD, removing Vicent-Gil et al. (2018) from the CPT RT analysis changed the overall result from non-significant to significant, and removing either sample from Halvorsen et al. (2012) from the simple-RT analysis changed the overall result from non-significant to significant.

3.3.8. Risk of bias

3.3.8.1. Risk of bias within studies

We rated whether the primary aim of each study may have influenced publication bias by grouping studies based on their aim. Out of the 103 included studies, eleven studies investigated PS or attention specifically as their primary aim (*k*=3 PS, *k*=7 attention, and *k*=1 both PS and attention; *k*=4 MDD, *k*=6 BD, *k*=1 both MDD and BD). We considered these studies to be the most susceptible to publication bias, given the potential motivation to publish significant results for the primary measures of PS and/or attention. Eighty-eight studies investigated PS and/or attention as part of a wider neuropsychological battery, with a primary aim to assess cognitive function more generally. We considered these studies moderately susceptible to publication bias since it may be of less relevance to report deficits in PS and/or attention in these studies, but there perhaps existed some motivation to present cognitive deficits generally. Four studies reported neuropsychological/cognitive scores secondary to another primary aim; we considered these as the least susceptible to publication bias. Any studies with a primary to assess PS or attention were removed from each comparison in the meta-analysis as a sensitivity analysis to test their effect on the overall result; none of the results were affected.

3.3.8.2. Risk of bias across studies

Funnel plots and Egger's test of publication bias were used to assess reporting bias for comparisons with at least ten studies. Only eleven comparisons had ten or more studies. For BD, these were: DSST (number correct), TMT-A (time to complete), verbal fluency, phonemic fluency, CPT d', and CPT RT. For MDD, these were: DSST (number correct), TMT-A (time to complete), choice-RT, verbal fluency, and Stroop simple trials. Visual inspection of funnel plots and Egger's test of publication bias suggested that no comparisons showed evidence of being influenced by publication bias (all *p*s>.05). A summary of the results of Egger's test are presented in Table 23 in Appendix A and the funnel plot for each comparison are presented in Appendix A (Figure 50 to Figure 60). From these results, along with the results of the sensitivity analysis reported above (removing studies with a primary aim to measure PS or attention), we considered the risk of publication bias to be minimal.

3.3.8.3. Risk of bias of this review

There are a few sources of potential bias in our methods. Firstly, due to the large number of search results from our initial search terms, we applied some limits to our search to reduce the search results to a manageable number. This involved excluding studies with "ADHD", "dementia" or "CBT" in the title. This may have excluded studies that contained a relevant sample of BD or MDD patients, for example, studies that compared neuropsychological performance of mood disorder groups to ADHD or dementia patients and a control group. Therefore, there may be sources of data that we missed. Similarly, we excluded studies when it was unclear if the study met our inclusion criteria. For example, our inclusion criteria stated that mood disorder samples must have no comorbidities, and in some cases, studies did not provide information on comorbidity in the publications. We contacted authors to request this information, but if we did not receive a response from authors and/or the details could not be confirmed, we excluded the study. This may have excluded potentially appropriate samples that would have been included if the appropriate information could be accessed. However, despite these exclusions, a large number of studies were included in the final sample, so we expect that missing these studies would not have changed the overall results considerably.

3.4. DISCUSSION

This review aimed to investigate the extent to which PS and SA are impaired in BD and MDD. We aimed to assess the magnitude of impairment for individual neuropsychological outcome measures. To our knowledge, this was the first meta-analysis to focus on PS and SA in mood disorders. The overall result suggests that BD and MDD are both associated with impairments in PS and SA compared to controls. An impairment was found across most tests of PS and SA.

3.4.1. BD results summary and comparison to previous research

Our results are in line with other meta-analyses of neuropsychological functioning in BD, showing impairments on tests of PS and SA in euthymic and symptomatic BD compared to controls, with moderate to large ESs (Bortolato et al., 2015; Kurtz & Gerraty, 2009; Torres et al., 2007). Large ESs were found for DSST, Stroop simple trials, and PS composite scores; the remaining impairments were of moderate ES. This suggests that some tests may be particularly sensitive to slower processing in BD. Previous research similarly showed that large ESs are found for tests of PS (Bo et al., 2017). Out of all the PS measures, only TMT-A number

of errors was not significantly impaired. SA was also impaired across most outcome measures. CPT RT and CPT RT variability were not impaired, which contradicts previous meta-analyses (Torres et al., 2007). This could suggest that impairments in SA may only be found when using accuracy or sensitivity measures, rather than reaction time. However, CPT RT was impaired in the manic subgroup, and CPT RT variability was impaired in the paper with a depressed sample (Gallagher, Nilsson, et al., 2015), suggesting this feature of SA may be a state-related impairment. Alternatively, this discrepancy may be explained by our exclusion criteria, as we removed studies that contained samples with comorbidities, whereas previous meta-analyses did not (Torres et al., 2007).

Where possible, we performed subgroup analysis to assess PS and SA in different mood states. Most of the outcome measures did not show significant subgroup differences, suggesting that impairments in PS and SA may be trait-features of the disorder that are present in euthymia. However, euthymic subgroups did not show impairments on some of the CPT RT measures, as mentioned above, so some abilities may be dependent on symptomatic states. Data from patients experiencing mania were only available from three studies (Mahlberg et al., 2008; Sweeney et al., 2000; Yadav et al., 2011) and there were sufficient data to meta-analyse TMT-A, simple-RT and choice-RT tasks. Subgroup analyses suggested that mania was associated with worse performance on TMT-A (time to complete) than other patient subgroups. These results are in line with previous research that suggests a more severe impairment in PS in manic BD patients (Kurtz & Gerraty, 2009).

Kurtz and Gerraty (2009) found that cognitive impairment is more severe in symptomatic states but that the effect of mood may depend on the instrument used to measure cognitive functioning. Our results are consistent with this, as only some of the tests we used to measure PS and SA showed differences between symptomatic and euthymic states. However, it should be noted that there were insufficient data to perform subgroup analyses for all neuropsychological outcome measures. For some comparisons (TMT-A, PS composite score, and CPT RT), subgroup differences suggested that the samples that were not specified by mood state were significantly different from the other subgroups. It may be the case that grouping participants across mood states can confound effects and conceal impairments associated with particular mood states.

3.4.2. MDD results summary and comparison to previous research

Patients with MDD had impairments in PS and SA compared to healthy controls with moderate to large ESs, which is in line with previous meta-analyses (Parkinson et al., 2020; X. Wang et al., 2020). We found large ESs for CPT RT variability (iSD) and PS composite score; the remaining comparisons had moderate ESs. These ESs were similar in magnitude to those reported by Wang et al. (2020). Similar ESs were found for category fluency and verbal fluency, which follows on from a previous meta-analysis that found similar ESs in both types of verbal fluency tasks after controlling for clinical characteristics (Henry & Crawford, 2005). Only simple-RT tasks and CPT RT CoV were not significantly impaired overall. The non-significant result for simple-RT tasks appeared to be driven by one sample of remitted patients, while the depressed subgroup showed a significant impairment. The result for CPT RT CoV was trending towards significance and the other measure of RT variability (iSD) suggested that variability of RT is increased in MDD compared to controls with a large ES.

Most of the outcome measures did not show significant subgroup differences, suggesting that poor PS and SA may be trait-features of MDD that can be found in remission. However, remitted subgroups did not show impairments on DSST, RT tasks, and Stroop simple trials, suggesting some impairments may be only present in acute phases, or that only certain neuropsychological tests are sensitive to this impairment. A previous meta-analysis similarly found a normalisation of the impairment on DSST, TMT-A, and PS composite in remission (Ahern & Semkovska, 2017), however they only meta-analysed data from 2 studies for each of these comparisons. Other meta-analyses found that remitted MDD patients showed impairments on most tests of PS, including DSST, TMT-A, simple-RT, and verbal fluency (Bora et al., 2013; Semkovska et al., 2019). The review by Semkovska et al. (2019) contained more studies for each comparison than our review, perhaps due to differences in inclusion criteria, so it may be the case that our null findings were a result of a smaller dataset. In some cases, the subgroup difference was driven by the sample with unspecified or mixed mood states (i.e., for TMT-A time and PS composite scores), suggesting that it may be the case that combining patients in different mood states may confound results for some comparisons.

3.4.3. Heterogeneity and possible confounds

While we found an overall impairment in PS and SA in mood disorders, the heterogeneity of cognitive functioning in patients remains an important factor that should not be overlooked.

Since previous research reported heterogeneity due to the neuropsychological instrument used (Tsitsipa & Fountoulakis, 2015), we attempted to mitigate this confound in our study by separating meta-analyses by neuropsychological outcome variable. Despite this, there was still moderate to high heterogeneity for most comparisons, even within mood state subgroups, similar to previous meta-analyses (Bo et al., 2017; Cullen et al., 2016; Parkinson et al., 2020; X. Wang et al., 2020). Heterogeneity may arise from many sources, including clinical and demographic confounds and family history. Since we created subgroups based on mood state, there were insufficient data to further sub-divide the data to investigate other sources of heterogeneity. These potential sources of heterogeneity should be considered in interpreting our results and in future research. It should also be noted that the *Q* statistic estimates the degree of heterogeneity between studies but cannot measure the degree of heterogeneity within each study, so potential sources of heterogeneity that contribute to a study mean should be considered.

Age may confound neuropsychological performance, with the effect being apparent in older adults (e.g., when comparing participants under and over 60 years old; Lim et al., 2013). We excluded studies with participants younger than 18 and older than 65 years old, so we do not expect that age confounded our results. We did not control for other demographic factors such age gender and education, however, these have been shown to have little effect on cognitive impairment in patients (Semkovska et al., 2019). Clinical characteristics such as earlier age of onset of the disorder and longer illness duration have been shown to be associated with more severe impairment in PS and SA in mood disorders (Bora et al., 2009; Bortolato et al., 2015; Cullen et al., 2016; Mann-Wrobel et al., 2011). Semkovska et al. (2019) assessed moderator effects in their meta-analysis of neuropsychological functioning in remitted MDD and found that the number of previous depressive episodes explained the largest relative variance in ESs. Other studies similarly found that number of episodes (Kriesche et al., 2022) and illness severity (McDermott & Ebmeier, 2009) are related to cognitive performance, including PS. The presence of psychosis and number of manic episodes have also been related to more severe cognitive dysfunction (Bora, 2018; Bora et al., 2010; Bourne et al., 2013). Other research found that cognitive impairment in BD was not explained by history of psychosis or number of episodes (Demmo et al., 2016). Subtype of BD has also been considered as a potential source of heterogeneity, with mixed findings: some studies suggest that BD-I and BD-II show distinctive neuropsychological of features (Bora, 2018), while

others do not report a difference in cognitive impairment between the two sub-groups (Cullen et al., 2016; Tsitsipa & Fountoulakis, 2015). Medication has likewise been related to cognitive impairment in some studies (Bora et al., 2009; Cullen et al., 2016; Lim et al., 2013), however, others found that this was not the case in euthymic BD (Bourne et al., 2013; Goswami et al., 2009). Few of the studies included in our analysis had a sample of completely medicated or completed unmedicated patients, or a whole sample of psychotic, or non-psychotic patients, so we did not investigate these effects. However, the heterogeneity in our results may have been driven by differences in many clinical characteristics and family history.

Another source of heterogeneity is the presence of childhood trauma and adversity: people with mood disorders have higher rates of childhood trauma than healthy controls (Mandelli et al., 2015; Watson et al., 2014), and childhood trauma has been associated with cognitive impairment in mood disorders (Bücker et al., 2013; Jørgensen et al., 2023). Therefore, high levels of trauma and adversity in the samples may have introduced heterogeneity in cognitive performance. Similarly, the presence of Autism Spectrum Disorders (ASD) in patient groups may confound results: mood disorders and ASD are often comorbid (Joshi et al., 2013; Leyfer et al., 2006) and ASD is also associated with cognitive impairment (Eack et al., 2013). Research suggests that children with comorbid ASD and BD show poorer attention and EF than children with a single diagnosis or either ASD or BD alone (A. S. Weissman & Bates, 2010), therefore the presence of people with comorbid ASD in BD samples likely adds to the heterogeneity in cognitive scores. ASD is rarely screened for in studies of cognitive functioning in mood disorders, therefore, this may contribute to heterogeneity in the literature. Alongside the lack of studies that screen for ASD, researchers have highlighted a lack of appropriate tools to better tools to assess mood disorders in ASD (Oakley et al., 2021). Future studies should therefore develop better tools to assess mood disorders in ASD and screen for ASD in mood disorder groups.

3.4.4. Limitations of the review

The present review has some limitations. Meta-analysis is prone to over-inclusion of studies with positive results (Dickersin, 1997; Higgins, Thomas, et al., 2022). We assessed the risk of publication bias in our study as low, however, we did not further assess risk of bias within studies formally. Studies that may have been particularly vulnerable to bias may therefore have skewed our results. Previous meta-analyses on cognitive impairments in mood disorders

found that risk of bias was unclear overall, as many studies do not provide the information needed to complete risk of bias analysis (Kriesche et al., 2022). Subgroup analyses were conducted where we had enough data, however, these are observational analyses which are prone to false positive and negatives, especially when multiple subgroup analyses are conducted (Higgins, Thomas, et al., 2022). Although we specified the intention to perform these subgroup analyses a priori, the comparisons between mood states should be interpreted with caution and should be retested in future research. We were not able to perform subgroup analyses for many of the comparisons due to lack of studies.

There were several steps in our method that may have limited the number of search results. Due to a large number of search results in our initial search (*n*=294,565), we applied some post-hoc limits to our search strategy: we excluded studies with "ADHD", "Cognitive behavioural therapy" or "dementia" in their titles. These topics were chosen as the most common irrelevant topics after a visual inspection of the initial search results. Studies on these topics likely included samples that were not relevant for our review (e.g., children for ADHD and older adults for dementia studies), however, this may have concurrently excluded suitable samples. We also excluded studies with patients with comorbidities. On one hand, removing the confounding effects of other diagnoses may be a strength, but on the other hand, this limited the number of studies in our results and may have occluded relevant datasets.

3.4.5. Neuropsychological methodology

Core cognitive functions are extremely difficult to measure in isolation and most neuropsychological tests will naturally capture other abilities as well. We included tests that are commonly used to measure PS and SA in the literature, however, there is not a consensus on which specific cognitive function(s) are being measured by each test and their validity should be considered. Previous systematic reviews found that individual neuropsychological tests are often considered to represent different cognitive functions across the literature (Cardenas et al., 2016). For example, DSST is commonly considered a measure of PS, however it was originally designed to measure general cognitive functioning and is sometimes considered to also test elements of WM and attention (Joy et al., 2000; Joy, Fein, et al., 2003). TMT-A is sometimes considered a test of attention (X. Wang et al., 2020) and verbal fluency is sometimes considered part of EF or considered as a separate function by itself (Tsitsipa & Fountoulakis, 2015). However, some research suggested that fluency better represents

deficits in PS in depression (Henry & Crawford, 2005). Similarly, different components of the CPT have been used for different purposes, e.g., RT has been used to measure PS (Thompson et al., 2005). While our meta-analysis provides evidence for impairments in PS and SA in mood disorders, we note that our results may also have implications for other impairments in cognitive functions such as general attention, WM, and EF.

The choice of outcome measure appears to affect the outcome of the result. BD and MDD took longer to complete TMT-A tasks than controls, however no impairment was found when using number of errors. This suggest that there may be elements of tasks such as time-based elements, that are more sensitive to PS impairments in mood disorder groups than accuracy measures. However, CPT RTs were not significantly longer in BD compared to controls, whereas accuracy measures were worse in patients, suggesting accuracy of CPTs may be a more sensitive measure of SA deficits in BD than RT measures. Similarly, simple-RT tests did not distinguish MDD patients from controls, but choice-RT tasks did. Patients performed worse than controls on traditional graphomotor tasks as well as computerised tasks.

Few studies included in our review attempted to apply statistical models other than the mean to measure PS and SA more precisely. Only five of the included studies reported variability in RTs from CPTs (Gallagher, Nilsson, et al., 2015; D. Kim et al., 2015; Naim-Feil et al., 2016; Schmidt et al., 2021; Yamamoto & Shimada, 2012). Findings on RT variability are limited but suggest that RT is more variable in depressed states compared to euthymia. Only one study used statistical modelling of RT data to assess cognitive functions: Gallagher et al. (2015) utilised ex-Gaussian modelling to get three parameters that reflect different components of the RT data. One of these components, *tau*, was sensitive to cognitive impairment in BD in euthymia and in depressed states. There was no difference in *tau* between MDD and controls, but there was a trend towards higher *tau* in MDD. Further research should apply these modelling methods to RT data to replicate these findings and test whether such statistical models can help to disentangle separate cognitive functions.

3.4.6. Conclusions and implications

Our review provided an updated overview of the nature of core cognitive impairment in mood disorders and provided insights and considerations for future research. BD and MDD were associated with slower PS and poorer SA compared to controls, which was evident in most neuropsychological tests we reviewed. Impairments were found in both euthymic/remitted

and symptomatic states. However, the presence of impairment in remitted states varied, with some comparisons showing impairment but to a lesser degree than in acute states and some cases showing no impairment in euthymia. There is a lack of data for manic and depressed groups. The magnitude of impairment in BD or MDD appears to depend on the instrument used, which suggests that researchers should carefully consider the design of their neuropsychological measures. More research is needed to deduce the specific cognitive functions that each instrument tests and to attempt to disentangle specific cognitive functions from scores. Statistical models should be employed to help deconstruct neuropsychological scores and RTs into separate cognitive components. Given that we have established core cognitive impairments in BD and MDD, the next step is to assess the impact of these impairments on wider cognitive functioning.

Chapter 4 COGNITIVE HIERARCHY IN MOOD DISORDERS

4.1. OVERVIEW

In Chapter 3, BD and MDD groups showed impairments in PS and SA that were apparent across mood states and with most neuropsychological instruments. The literature also supports the hypothesis that EF is widely impaired in mood disorders (Cotrena et al., 2020; Cullen et al., 2016; Semkovska et al., 2019). As described in section 1.3, PS, SA, and EF may have a role as 'core' cognitive functions that influence wider cognitive functioning and, when impaired as a consequence of psychopathology, may have a knock-on effect causing secondary general cognitive dysfunction. The concept of a cognitive hierarchy is important to examine because if a perceived memory deficit, for example, is in fact caused by slower processing, poor attention, or executive dysfunction, this would have implications for our understanding of cognition in mood disorders, as well as for treatments and interventions.

4.1.1. Cognitive hierarchy in BD and MDD

There is limited research on the hierarchical nature of cognitive dysfunction in mood disorders. Chapter 3 showed that PS and SA are impaired in patients and studies suggest that impairments in PS, attention, and EF may play a role in wider cognitive dysfunction (Daban et al., 2012; Gallagher et al., 2014; Gallagher, Nilsson, et al., 2015). PS has been found to account for a large part of the variance in WM, verbal memory, and visuo-spatial functioning in people with BD and FDRs (Antila et al., 2011; Kieseppä et al., 2005). Another study demonstrated that PS and EF explain memory impairments in BD (Thompson et al., 2009). PS and EF also seem to be primary deficits in MDD (J. Liu et al., 2019), and PS appears to mediate the relationship between depression status and verbal, visuo-spatial, and WM impairments in MDD (Zaremba et al., 2019). In another study, executive dysfunction in MDD was explained by an attentional deficit (Nilsson et al., 2016). These studies suggest that PS, EF, and attention may indeed play a role in wider cognitive dysfunction in both BD and MDD. However, no studies to our knowledge have assessed the role of SA. SA may be of particular interest, rather than attention more generally, since SA, as conceptualised by ex-Gaussian models of RT data, appears to be sensitive to cognitive impairment in mood disorders (Gallagher, Nilsson, et al., 2015; Nilsson et al., 2014). More research is needed to uncover the nature of the relationships between

cognitive functions and establish whether primary impairments in PS, SA, and EF can explain secondary cognitive dysfunction.

4.1.2. Effects of diagnostic group and mood state

BD and MDD may both involve a hierarchy of cognitive dysfunction, however, whether this hierarchy is of a different nature in each diagnostic group is not known. It is unclear whether the same profile of cognitive impairments exist in BD and MDD, with some studies suggesting no difference between the groups (R. Lee et al., 2015; Porter et al., 2015), some suggesting a greater magnitude of impairment in BD compared to MDD (Zazula et al., 2021), and others suggesting distinct patterns of impairments (Tavares et al., 2007). A recent study utilised network graphs to assess the cognitive profile of BD and MDD, and while no group differences in any cognitive domain were found, networks highlighted different roles of memory and EF in each disorder (Galimberti et al., 2020). MDD appeared to have a more densely interconnected network than BD, with memory as the most central domain in the network, whereas EF was more central in BD. It may therefore be important to investigate the cognitive profile of BD and MDD separately. Interestingly, attention was found to moderate cognitive function in both groups in this study, however, this study was limited to the cognitive domains from the Montreal Cognitive Assessment battery, which did not include PS or SA. Mood state is also an important consideration in researching cognitive dysfunction: deficits may persist in euthymia or remission, perhaps to a lesser degree (Bourne et al., 2013; Kurtz & Gerraty, 2009). Other studies suggest impairments in core functions such as PS exist in symptomatic states but not in remission (Langenecker et al., 2010; Zaremba et al., 2019). The nature of a cognitive hierarchy may therefore vary between mood states and diagnostic groups.

4.1.3. Open questions in the literature, aims, and hypothesis

Overall, existing research suggests that slow processing, poor attention, and executive dysfunction may explain wider cognitive impairment in mood disorder groups, however, there is limited research that specifically tests the hierarchical nature of cognitive dysfunction. Previous studies tended to focus on PS or EF and have not yet tested the role of SA. Research should also account for relationships between core cognitive functions, however, no studies to the author's knowledge have tested whether PS, SA, and EF together can account for wider cognitive dysfunction in mood disorders. Investigating the nature of this cognitive hierarchy is

vital for unravelling the mechanisms underpinning cognitive dysfunction in BD and could help inform cognitive remediation interventions.

We therefore set out to extend current research on the presence of a hierarchy of cognitive impairment in mood disorders. We aimed to investigate the role of PS, SA, and EF on memory impairments in BD and MDD in different mood states, while controlling for potential confounds. We hypothesised that PS, SA, and EF would explain the between-group difference in memory between mood disorder patients and healthy controls.

4.2. METHODS

To assess the presence of cognitive hierarchy across different patient groups and mood states, an ideal study design would recruit different patient groups, for example, samples of BD and MDD patients in both euthymia/remission, depression, and mania (for BD), and a healthy control group, that were all matched on important demographic confounds such as age, sex, years of education, and premorbid IQ. The International Society for Bipolar Disorders (ISBD) Targeting Cognition Task Force have published recommendations for measuring cognitive functioning in BD, including using the following tests: CPTs, symbol coding tests, TMT-A, category fluency, verbal learning tasks, and tests of EF such as the Stroop test and TMT-B (Miskowiak, Burdick, et al., 2017; Yatham et al., 2010). The ISBD Targeting Cognition Task Force also recommended combining multiple test scores into one composite cognitive score, which may be a more robust outcome than a single cognition test (Miskowiak, Burdick, et al., 2017). Since our aim here was to investigate the role of separate cognitive domains, we aimed to create composite scores for each cognitive domain rather than one general cognitive composite score. In line with these recommendations, a comprehensive neuropsychological battery for this study would include several measures of core cognitive functions, such as: DSST, TMT-A and simple-RT for PS; CPT for SA; and digit span backwards, TMT-B, the Stroop task, and verbal fluency tasks for EF. Immediate and delayed scores from RAVLT could measure verbal memory, as well as several tasks to measure visuo-spatial memory, such as spatial recognition, pattern recognition, and spatial span tasks.

Methodology should be consistent across patient groups and the choice of neuropsychological tests should be informed by the ability to statistically disentangle cognitive processes from scores to improve task validity. For example, in the DSST, the symbol copy variant should be

run as well as the original variant, to subtract motor speed to better assess cognitive speed. For measures of EF, scores from simple parts of the tasks (e.g., TMT-A, Stroop simple trials, and digit span forwards) should be subtracted from more complex parts of the tasks (e.g., TMT-B score, Stroop colour-word trials, and digit span backwards, respectively) to remove speeded and attentional processes from the scores, theoretically leaving only EF components. Similarly, ex-Gaussian models should be applied to CPT data to disentangle elements of PS from SA. Due to external factors, data collection was not possible for this thesis, therefore, we chose extant datasets that matched this design as closely as possible.

Data from three previous studies that investigated neuropsychological performance in mood disorder patients and HCs were used. The first dataset was a sample of euthymic BD patients (BD-e; *n*=63; Thompson et al., 2005); the second was a sample of depressed BD patients (BD-d; *n*=42; Watson et al., 2012); and the third was a sample of patients with Major Depressive Disorder (MDD; *n*=41) who were currently depressed and unmedicated at the time of testing (Porter et al., 2003). The datasets differed on some potential confounds, such as mood state, medication, and neuropsychological assessment, so we analysed the datasets separately and interpreted the results for each diagnostic group and mood state.

4.2.1. Participants

Table 4 provides an overview of the datasets and details of each sample. Patients were recruited from outpatient centres in the Northeast of England; some BD-d patients were recruited from Christchurch, New Zealand. Diagnoses were made according to DSM–IV criteria (American Psychiatric Association, 1994) and confirmed using the Structured Clinical Interview for DSM–IV (SCID; First et al., 1995). BD-e and MDD had no psychiatric comorbidities; the BD-d dataset did not include data on comorbidities. Each study had a matched HC group without current or past psychiatric illness. Controls were recruited from the community by local advertisement and matched to each patient group on age and sex, as well as on other demographic variables for BD-e and MDD (see Table 4). All participants were 18-65 years old. Ethical approval was granted for each study by the local University and NHS ethics committees and participants provided informed consent.

Dataset	BD-e	BD-d	MDD
Diagnosis	Bipolar Disorder	Bipolar Disorder	Major Depressive Disorder
Mood state	Euthymic	Depressed	Depressed
Original study	Thompson (2005)	Watson (2012)	Porter (2003)
N patients	63	43	41
N controls	63	38	41
Mood state criteria	Euthymia prospectively defined by HRSD and YMRS scores ≤7 at assessment and after 1 month.	Depressive episode confirmed with SCID. Severity of mood symptoms assessed using HRSD-17 and YMRS.	Depressive confirmed with SCID. Severity of depression assessed using MADRS, HRSD-17, and BDI.
Medication	All were stabilised on prophylactic medication at testing, except 3 patients who were not taking medication.	Patient medication was stable (unchanged for 4 weeks before participation).	All patients were free of any psychotropic medication for at least 6 weeks.
Other clinical features	54 had BD type I, 9 had BD type II, and 5 were rapid cycling.	-	Single episode or recurrent MDD.
Control group matched on	Age, sex, years of education, race, handedness, and premorbid IQ,	Age and sex.	Age, sex, years of education, premorbid IQ, season of testing, and phase of menstrual cycle.
Neuropsychological tests for each cognitive domain			
Processing speed	DSST (number correct) TMT-A (time)	DSST (number correct) SCOLP	DSST (number correct)
Sustained attention	Vigil CPT number of errors Vigil CPT ex-Gaussian *tau*	Vigil CPT number of errors Vigil CPT ex-Gaussian *tau*	Vigil CPT number of errors Vigil CPT ex-Gaussian *tau*
Executive function	Digit backwards span (minus forwards span) TMT-B (time minus A time) Stroop CW (correct)	Digit backwards span (minus forwards span) Verbal fluency	ToL perfect solutions ToL average excess
Verbal learning and memory	RAVLT A1-5 total RAVLT % retained at A6 RAVLT % retained at A7	RAVLT A1-5 total RAVLT % retained at A6 RAVLT % retained at A7	RAVLT A1-5 total RAVLT % retained at A6 RAVLT % retained at A7
Visuo-spatial memory	Spatial recognition Pattern recognition SWM (number errors) SWM strategy Spatial Span	Spatial recognition Pattern recognition SWM (number of errors) Spatial Span	Spatial recognition Pattern recognition SWM (number errors) SWM (strategy)

Table 4: A summary of three datasets used in this analysis. BD-e=bipolar disorder euthymic; BD-d=bipolar disorder depressed; MDD=major depressive disorder; HRSD=Hamilton Rating Scale of Depression; YMRS=Young Mania Rating Scale; MADRS=Montgomery-Asberg Depression Rating Scale; BDI=Beck Depression Inventory; SCID=Structured Clinical Interview for DSM–IV;; DSST=Digit Symbol Substitution Test; TMT=Trail-Making Test; SCOLP=Speed and Capacity of Language Processing; CPT=Continuous Performance Test; CW=colour-word; ToL=Tower of London; RAVLT=Ray Auditory Verbal Learning Test; SWM=Spatial Working Memory.

4.2.2. Procedure

Participants completed a neuropsychological test battery that was administered according to standard instructions and manual protocols. For BD-e and MDD, neuropsychological assessment was completed at 2pm to control for possible effects of diurnal variation in performance and lasted between 1.5 to 2 hours. For BD-d, neuropsychological performance was measured at several timepoints as part of a longitudinal trial; baseline neuropsychological data were used here. Further details of the participants and procedure can be found in the original articles (Porter et al., 2003; Thompson et al., 2005; Watson et al., 2012).

4.2.3. Materials

The National Adult Reading Test (NART; Nelson, 1982) was used to measure premorbid IQ. Individual neuropsychological tests were selected from the available data to represent cognitive domains of processing speed (PS), sustained attention (SA), executive function (EF), verbal learning and memory (VM), and visuo-spatial memory (VS). We extracted neuropsychological scores that were as similar as possible to the other datasets, however, datasets used different instruments for some cognitive functions, so some domain scores were comprised of different tests. A summary of the tests used to represent each cognitive domain in each dataset are summarised in Table 4.

4.2.3.1. Processing speed (PS)

Measures of PS included the Digit Symbol Substitution Test (DSST), the Trail-Making Test (TMT) part A, and the Speed and Capacity of Language Processing (SCOLP) test. The DSST is a graphomotor test of PS (Joy, Fein, et al., 2003; Lezak, 1995). Participants transcribed a unique geometric symbol with a corresponding Arabic number, which was shown in a key at the top of the page. The number of correct responses in 90 seconds was used as the outcome measure. TMT is a graphomotor test of PS, attention, and EF (Reitan, 1958). In part A of the task, participants connected a series of numbered circles spread on a page in ascending order. TMT-A therefore captured elements of motor speed, cognitive speed, and attention. Time to complete the task was used as the outcome measure. SCOLP tests the speed and efficiency of cognitive processing (Baddeley et al., 1992) and has two subtests: speed of comprehension test and spot-the-word test. The former subtest was used here, as it provides a measure of current information PS, whereas 'spot the word' measures premorbid intellectual functioning. In speed of comprehension, participants read a series of statements that reflected simple and

common-sense knowledge about the world and categorised each statement as true or false as quickly as possible. Scores were the total number of correct responses in two minutes.

4.2.3.2. Sustained attention (SA)

The Vigil Continuous Performance Test (CPT) was used as a computerised measure of SA (Cegalis & Bowlin, 1991; Psychological Corporation, 1998). Participants viewed a continuous stream of letters on screen and pressed a single response button when the letter 'K' appeared after the letter 'A'. Errors of omission and commission and reaction time were recorded. We performed ex-Gaussian analysis on the reaction time data to extract ex-Gaussian metrics. The total number of errors and ex-Gaussian *tau* were used as outcome measures.

4.2.3.3. Executive function (EF)

The three datasets differed on tests of EF. For BD-e and BD-D, Digit Span was used to test WM and EF. Participants had to remember and verbally recall a series of numbers that gradually increased in length (Lezak et al., 2004). In the 'forward span' subtest, participants recalled the series of numbers in the same order in which they heard them, measuring attention and immediate memory. In 'backwards span', participants recalled the numbers in reverse order, involving updating WM, a component of EF (Miyake, Friedman, et al., 2000). The score for each subtest is the maximum span attained. For the outcome measure, we subtracted forwards span from backwards span to subtract out the attentional component (Nilsson et al., 2016). For BD-e, part B of the TMT was also used to measure EF (Reitan, 1958). In TMT-B, participants connected a series of circles on the page that were either numbered or contained a letter, in alternate ascending and alphabetical order (e.g., 1-A-2-B-3-C, etc.). The time taken to complete the task was recorded. TMT-B is thought to reflect the shifting component of EF (Miyake, Friedman, et al., 2000), however, it also captures elements of PS and attention, so we subtracted TMT-A from TMT-B time to subtract out attentional and speeded components (Nilsson et al., 2016).

For BD-e, the Stroop colour-word test measured the inhibition component of EF (Miyake, Friedman, et al., 2000; Stroop, 1935). The Stroop task includes a colour naming block, a word reading block, and a colour-word (CW) block. In CW, participants viewed words describing colours on a screen (e.g., 'red') with the text of the words presented in various font colours. Participants said aloud the colour of the font whilst ignoring the word itself, requiring inhibition of the automatic response to read the word. In the first two blocks, participants

either named the colour of a block of colour presented on screen, or read a word presented in black font, therefore these blocks reflected automatic processes of PS and attention. Scores were the number of correct responses. We subtracted colour naming scores from CW scores to subtract out the attentional and speeded components (Nilsson et al., 2016).

For BD-d, verbal fluency was used to measure EF alongside digit span. Subjects named aloud as many words possible beginning with F, A, or S (phonemic fluency; Lezak et al., 2004). Scores were the number of correct words in 60 seconds. The MDD dataset did not share any tests of EF with the other two datasets, so the Tower of London (ToL) task measured EF (Humes et al., 1997). In ToL, participants rearranged a set of discs to match a target order in the fewest number of moves possible. The average number of perfect solutions and the average number of excess moves were used as outcome variables.

4.2.3.4. Verbal learning and memory (VM)

Several measures from the Rey Auditory Verbal Learning Test (RAVLT; Rey, 1964) tested VM. Participants heard a list of 15 words (list A) five times and had to verbally recall the words after each trial (trials A1 to A5). A different list of words (list B) was then read aloud, and participants recalled this list. Participants were then asked to recall list A without hearing list A again (trial A6). After a 30-minute delay, participants recalled list A again (trial A7). Total words remembered from trials A1-A5 represented immediate verbal learning, A6 represented immediate recall, and A7 represented delayed recall. Since performance on A6 and A7 depended on initial verbal learning, these scores were calculated as a percentage of the maximum score from trials A1-5 to better reflect verbal memory rather than learning (Nilsson et al., 2016; Thompson et al., 2009).

4.2.3.5. Visuo-spatial memory (VS)

Several computerised tasks from the CANTAB (Cambridge Cognition, 2015) were used to represent VS. In Spatial Recognition, a white square appeared at various locations in sequence on a screen and participants had to recognise which location they had seen before out of two choices. In Pattern Recognition, participants saw a series of visual patterns on screen then had to choose between a pattern they had already seen and a novel pattern. Another recognition test was conducted after a delay of 10-20 minutes. For both tasks, the outcome measure was the percentage of correct trials. A Spatial Working Memory (SMW) task was also used to measure VS: several coloured squares were presented on screen and participants had to find

yellow tokens in some squares by using a process of elimination. Task difficulty increased by gradually increasing the number of squares in each trial up to twelve. Two outcome measures were used, where possible: number of errors (selecting squares that had already been selected; score collapsed over levels 4, 6 and 8) and strategy count (for sets of length 6 and 9; a high score represented a lower use of strategy). Finally, in Spatial Span, participants viewed white squares on a screen with some squares changing in colour in a sequence. Participants selected the squares that changed colour either in the same order that they were presented (forward span) or in reverse order (backwards span). The length of the sequence increased by one each trial, starting at two squares and ending at nine squares. The longest sequence successfully recalled (span length) was the outcome measure.

4.2.4. Data analysis

Data pre-processing and analysis was performed in R Studio version 4.1.2 (R Core Team, 2021). The same pre-processing pipeline was applied to each dataset separately.

4.2.4.1. Pre-processing

One BD-e patient had a substantial amount of missing neuropsychological data so were excluded from the analysis. Of the remaining participants, some had missing data for NART and neuropsychological variables: 1% of data were missing in the BD-e and BD-d datasets, and 3% of data were missing overall in the MDD dataset. For MDD, ToL outcome measures and Vigil *tau* had a relatively high portion of missing data compared to the other variables (9.8% and 15.8%, respectively). However, for each dataset, data were missing at random (i.e., not influenced by another variable) and was therefore considered suitable for imputation. Details of missing data are presented in Table 24 in Appendix B.

We imputed missing data for these variables using regression techniques, rather than conducting listwise deletion, pairwise deletion, or mean substitution, in order to retain sample size, preserve deviations from the mean, and retain the shape of the distribution. Also, regression techniques are more accurate when higher inter-correlations between the variables are expected, as expected in cognitive data (Raymond, 1986; Roth, 1994). Multiple Imputation by Chained Equations (MICE) is generally recommended when imputing missing data from multiple variables as it accounts for the statistical uncertainty in the imputations (Azur et al., 2011; Schafer & Graham, 2002). In MICE, missing data is imputed multiple times based on the observed values for other participants, creating multiple complete datasets.

Each of these datasets are then used in the planned analysis, and the results are pooled to give a final estimate of the results. However, in our pre-processing pipeline, several steps followed missing data imputation (creation of composite domain scores and regressing out confounding variables) which meant that MICE could not be applied in this way and a single imputed dataset was required to feed into the next step of pre-processing. Imputation using linear regression has been suggested to be an appropriate method when 1-15% of data is missing at random, where missing cognitive data can be considered missing at random if missing cases do not depend on cognitive ability (Raymond & Roberts, 1987; Roth, 1994). The *mice* package in R (van Buuren & Groothuis-Oudshoorn, 2011) was used to produce one imputed dataset where each missing data point was predicted using linear regression with age, education, NART, and all other neuropsychological scores as predictors. Stochastic regression was attempted, but led to implausible cognitive values, so linear regression was used. Imputation of missing data was done for patient and controls groups separately, since group differences in cognitive data were expected. Scatter plots of imputed data were checked and visually compared to the original data to detect whether any imputed values were extreme or unusual; imputed data were considered unbiased.

Cognitive variables were then standardised (z-scored) based on the control group mean and standard deviation (SD). Z-scores were reversed where appropriate so that a higher score reflected better performance. Composite scores for each cognitive domain were then calculated by averaging z-scores of the relevant neuropsychological variables. We controlled for age and pre-morbid IQ by regressing these variables on each cognitive domain and using the residuals from those regressions as the domain scores. Age and NART IQ were regressed out of the control group data using robust regression for each cognitive domain individually. To ensure that the residuals of these regressions were not skewed, outliers in the control group were Winsorized to 3 SDs above or below the mean to retain sample size and information about rank (Osborne, 2013). Box Cox transformations were then applied to the domain scores before the regressions (Box & Cox, 1964); details of the transformations performed for each composite domain score in each dataset are presented in Table 25 in Appendix B. Robust regression was used to mitigate issues with skewed distributions and outliers. The residuals following the age and NART regression for each domain variable were normally distributed in the HC group. The same Box Cox transformations that were applied to

the control group were then applied to the patient group, before regressing out age and NART IQ in the patient group using the regression model estimated for the control group.

4.2.4.2. Statistical analysis

To test group differences in demographics, clinical characteristics and cognitive performance, independent samples t-tests were conducted on the raw data (i.e., the untransformed data with missing data). Non-parametric Mann-Whitney U-tests were used for skewed data. Spearman correlations were used to test relationships between cognitive domains. Linear multiple regression models were used in the mediation and hierarchical regression analysis, since data were not extremely skewed, since regression is robust with respect to violations of the normality assumption, and to avoid over-transforming the data (Ernst & Albers, 2017).

4.2.4.3. Network graphs

Testing correlations between cognitive domains was an exploratory step, so we did not control for multiple comparisons. However, we used network graphs to illustrate relationships between cognitive domains and to serve as a more conservative estimate of relationships between variables. Network graphs were produced using the *qgraph* and *bootnet* packages in R (Epskamp et al., 2012, 2018). Networks were constructed using Gaussian Markov random field estimations estimated using a Least Absolute Shrinkage and Selection Operator (LASSO) penalty. A tuning parameter, lambda, controlled the level of sparsity of the network. Lambda was selected automatically using an Extended Bayesian Information Criterion (EBIC) model selection. This regularisation method causes small connections to automatically shrink to 0 to yield a parsimonious network (Epskamp & Fried, 2018). A hyperparameter for EBIC, gamma, was set manually to 0 and controlled the degree to which simpler models are preferred. Gamma can range from 0 to 0.5, where at higher values, simpler models are preferred, resulting in fewer edges (connections) being retained. While gamma=0 (i.e., the standard Bayesian Information Criterion [BIC]) errs on the side of discovery, the network will still be sparser than a partial correlation network without any regularisation. LASSO regularisation with EBIC model selection has been shown to have high specificity (i.e., less false positives), but varying sensitivity (i.e., how well it can detect true edges in the network) (Epskamp & Fried, 2018). Therefore, an absence of an edge cannot be interpreted as there being no relationship in real life. The edges in the networks can be interpreted as partial correlations between the cognitive domains. The networks were used to illustrate potential relationships

and network metrics were not employed here as the small number of nodes in the network (*n*=5) would make it difficult to interpret network parameters.

4.2.4.4. Mediation analysis

For multiple and hierarchical regression analyses, we analysed the whole dataset (i.e., patients and controls together) and used group status a variable. We tested whether each core cognitive variable mediated the effect of group status on memory using the method outlined by Baron and Kenny (1986). According to this method, mediation is present when: the predictor significantly predicts the outcome; the mediator significantly predicts the outcome; the predictor significantly predicts the mediator; and the predictor no longer significantly predicts the outcome when both the predictor and mediator variables are entered into the regression. There was a basis for partial mediation when group status still had a significant effect on memory after adding a core cognitive domain to the model, but the *β*-value of group status appeared to be weakened. In this case, a Sobel test was used to test whether partial mediation was present (Sobel, 1982). A series of regressions were therefore performed: simple regression models tested whether group status (patient vs control) had a significant effect on memory, and whether each of the core cognitive variables had a significant effect on memory. Following this, multiple regression models tested whether each core cognitive domain could explain the group difference in memory. Separate models were run for each core cognitive domain score (PS, SA, or EF) as the mediator along with group status as the predictor and memory (VM or VS) as the outcome.

4.2.4.5. Hierarchical regression models

Hierarchical regression was used to test whether PS, SA, and EF can explain the group difference in VM and VS, while accounting for the other core cognitive variables. Hierarchical regression was chosen over other regression-based methods for several reasons: 1) it allows for comparison to previous analyses (Thompson et al., 2009); 2) the predictors and their order of entry were chosen based on theoretical grounds, 3) a predetermined hierarchical order of entry of predictors was used rather than step-wise methods to avoid risks of over- or under-fitting (Harrell, 2016). The three core cognitive domains and group status were the predictors and were added sequentially, with group included at each step to see if there was a significant effect of group on memory after accounting for core cognitive functions. The order of the variables was switched each time to test whether each core cognitive domain could account

for a significant proportion of additional variance after accounting for the other core cognitive variable(s). In model 1, PS and group were entered first, then SA was added in step 2, and EF was added in step 3. In model 2, SA and group were entered first, then PS was added in step 2, and EF was added in step 3. In model 3, EF and group were entered first, then PS was added in step 2, and SA was added in step 3. In comparing each step of the hierarchical regression models, *F*-change statistics were used. *F*-change statistics are also known as nested-model tests and determine the significance of an R^2 change, where a significant *F*-change suggests the added variable significantly improves the model prediction. Models were repeated for VM and VS separately as the dependent variable.

4.3. RESULTS

4.3.1. Group differences

Descriptive statistics and group differences in demographics, clinical characteristics and neuropsychological performance for BD-e, BD-d, and MDD are shown in Table 5, Table 6, and Table 7, respectively. Group differences in clinical characteristics and neuropsychological performance have been reported previously (Douglas et al., 2018; Gallagher et al., 2015; Moss et al., 2016; Thompson et al., 2005; Thompson et al., 2009) but will be reported here since our samples differed slightly from the original publications, for example, where we removed participants with a substantial amount of missing neuropsychological data. BD-e and BD-d patients did not differ from matched HCs on sex, age, years of education, or premorbid IQ, but scored significantly higher than HCs on the depression and mania scales. MDD patients had fewer years of education than HCs, but did not differ in sex, age, or premorbid IQ. HCs matched to the MDD sample did not complete mood scales, so depression scores could not be compared between groups. We correlated potential confounding variables to assess relationships between them for each group; details can be found in Appendix B. NART IQ was significantly correlated with education in all patient and control groups (p<.05), except the MDD control group. Age was not related to NART IQ or education in any of the samples, except the MDD control group and BD-d control group, where age was correlated with NART.

	Healthy controls				Bipolar disorder euthymic				Group differences			
	N	**Mean [N]**	**SD [%]**	**Median**	**N**	**Mean [N]**	**SD [%]**	**Median**	**Statistic (*t*/*w*)**	***df***	***p***	**Cohen's *d***
Sex (N/% female)	63	[37]	[59]	-	63	[37]	[59]	-	χ^2=0	1	1.00	-
Age (years)	63	45.40	9.08	47.00	63	44.43	8.64	46.00	1827.0	-	.442	0.11
Education (years)	63	14.17	3.05	13.00	63	14.24	2.96	14.00	2051.0	-	.746	0.02
HAM-17	63	0.65	1.06	0.00	63	1.43	1.60	1.00	2633.0*	-	.001	0.58
Young Mania Rating Scale	63	0.30	0.78	0.00	63	0.81	1.54	0.00	2315.0*	-	.031	0.42
National Adult Reading Test	63	110.00	9.20	110.00	62	109.61	10.22	110.00	-0.22	121.2	.824	0.04
Processing Speed	**63**	**0.00**	**0.85**	**0.05**	**63**	**-0.85**	**1.24**	**-0.62**	**-4.49***	**110.1**	**<.001**	**0.81**
DSST (correct)	63	61.24	10.00	62.00	62	50.47	13.16	51.50	-5.15*	113.84	<.001	0.93
TMT-A (time)	63	32.06	8.83	30.66	62	37.41	13.09	34.81	2392.0*	-	.030	0.48
Sustained Attention	**63**	**0.00**	**0.89**	**0.17**	**63**	**-0.82**	**1.60**	**-0.40**	**1279.0***	**-**	**.001**	**0.64**
Vigil CPT total errors	63	3.43	4.12	2.00	62	7.89	9.51	5.00	2541.0*	-	.004	0.62
Vigil CPT ex-Gaussian *tau*	58	67.18	23.09	66.60	61	80.57	32.88	77.01	2208.0*	-	.020	0.47
Executive Function	**63**	**0.00**	**0.60**	**0.03**	**63**	**-0.39**	**0.76**	**-0.18**	**1394.0***	**-**	**.004**	**0.57**
Digit Span (backwards minus forwards)	63	-1.76	1.35	-2.00	62	-2.23	1.14	-2.00	1563.0*	-	.048	0.37
TMT-B (minus TMT-A time)	63	33.54	22.38	26.00	61	35.22	23.74	30.00	2063.0	-	.482	0.07
Stroop (CW minus colour correct)	62	-11.71	11.91	-5.50	61	-20.93	19.22	-17.00	1419.0*	-	.017	0.58
Verbal memory	**63**	**0.00**	**0.78**	**-0.04**	**63**	**-0.46**	**1.00**	**-0.30**	**1498.0***	**-**	**.018**	**0.52**
RAVLT A1 to 5 total	63	51.65	7.90	52.00	62	46.74	9.07	47.50	-3.22*	120.2	.002	0.58
RAVLT % retained at A6	63	0.84	0.13	0.85	62	0.79	0.17	0.80	1650.0	-	.134	0.34
RAVLT % retained at A7	63	0.82	0.14	0.83	62	0.77	0.18	0.79	1594.0	-	.076	0.34
Visuo-spatial memory	**63**	**0.00**	**0.65**	**-0.01**	**63**	**-0.53**	**0.89**	**-0.42**	**-3.84***	**113.5**	**<.001**	**0.69**
Spatial Recognition % correct	63	0.82	0.10	0.85	63	0.74	0.13	0.75	1272.0*	-	<.001	0.65
Pattern Recognition % correct	63	0.90	0.09	0.92	63	0.85	0.13	0.85	1604.0	-	.061	0.44
Spatial Span	63	5.92	1.20	6.00	61	5.30	1.30	5.00	1464.0*	-	.017	0.51
SWM total errors	63	29.06	17.71	28.00	63	40.57	24.20	45.00	2549.0*	-	.006	0.55
SWM strategy count	63	34.22	5.60	35.00	63	35.37	5.91	36.00	2351.0	-	.073	0.20

*Table 5: Demographics, clinical characteristics, neuropsychological scores, and group differences for euthymic bipolar patients and matched controls. *Significant at the 0.05 level. SD=standard deviation; df=degrees of freedom; HAM-17 Hamilton Depression Rating Scale 17-item; YMRS=Young Mania Rating Scale; DSST=Digit Symbol Substitution Test; TMT=Trail-Making Test; CW=Colour-Word; CPT=Continuous Performance Test; RAVLT=Rey Auditory Verbal Learning Test; SWM=Spatial Working Memory.*

	Healthy controls				Bipolar disorder depressed				Group differences			
	N	**Mean [N]**	**SD [%]**	**Median**	**N**	**Mean [N]**	**SD [%]**	**Median**	**Statistic (*t/w*)**	***df***	***p***	**Cohen's *d***
Sex (N/% female)	38	[19]	[45]	-	43	[16]	[37]	-	χ^2=.87	1	.350	-
Age (years)	38	44.42	13.71	46.50	43	47.84	10.22	49.00	915.5	-	.353	0.29
Education (years)	31	14.77	2.12	16.00	40	14.32	2.67	15.00	566.0	-	.528	0.19
HAM-17	29	0.48	1.27	0.00	40	19.45	5.18	19.00	1160.0*	-	<.001	4.77
YMRS	25	0.12	0.44	0.00	40	1.20	1.60	0.00	694.5*	-	.001	0.85
National Adult Reading Test	37	113.76	11.75	118.00	43	110.79	10.03	112.00	637.5	-	.128	0.28
Processing Speed	**38**	**0.00**	**0.83**	**-0.09**	**43**	**-1.05**	**0.84**	**-0.98**	**-5.60***	**78.0**	**<.001**	**1.26**
DSST (correct)	38	60.16	11.79	60.00	42	47.64	11.66	50.00	-4.77*	77.0	<.001	1.08
SCOLP (correct)	37	75.81	16.68	75.00	42	58.29	15.04	60.50	-4.88*	73.1	<.001	1.12
Sustained Attention	**38**	**0.00**	**0.88**	**0.11**	**43**	**-1.01**	**1.43**	**-0.71**	**431.0***	**-**	**<.001**	**0.85**
Vigil CPT total errors	38	5.16	7.39	2.00	43	11.33	10.94	9.00	1159.5*	-	.001	0.66
Vigil CPT ex-Gaussian *tau*	31	69.60	39.61	71.79	41	120.36	64.63	102.57	965.0*	-	<.001	0.93
Executive Function	**38**	**0.00**	**0.76**	**0.19**	**43**	**-0.02**	**0.65**	**-0.05**	**770.0**	**-**	**.660**	**0.03**
Digit Span (backwards minus forwards)	38	-2.21	2.40	-2.00	43	-1.33	2.15	-1.00	1.74	74.9	.086	0.4
Verbal fluency	38	43.47	10.80	45.00	43	39.02	8.75	39.00	-2.02*	71.3	.047	0.46
Verbal memory	**38**	**0.00**	**0.85**	**0.07**	**43**	**-0.36**	**0.84**	**-0.52**	**-1.92**	**77.5**	**.059**	**0.43**
RAVLT A1 to 5 total	38	47.50	8.70	47.50	43	41.56	8.80	40.00	-3.05*	78.0	.003	0.69
RAVLT % retained at A6	38	0.79	0.26	0.85	42	0.77	0.20	0.82	725.5	-	.487	0.06
RAVLT % retained at A7	38	0.73	0.23	0.77	42	0.66	0.27	0.67	-1.39	77.6	.169	0.31
Visuo-spatial memory	**38**	**0.00**	**0.80**	**0.10**	**43**	**-0.32**	**0.67**	**-0.29**	**-1.96**	**72.9**	**.054**	**0.45**
Spatial Recognition % correct	38	14.87	3.07	16.00	43	14.23	3.09	14.00	-0.93	77.9	.357	0.21
Pattern Recognition % correct	38	22.11	2.17	23.00	43	21.47	2.69	22.00	719.5	-	.349	0.26
Spatial Span	38	5.95	1.23	6.00	43	5.28	0.88	5.00	580.0*	-	.016	0.64
SWM total errors	38	23.71	21.65	17.00	43	29.05	18.57	31.00	960.5	-	.176	0.27

*Table 6: Demographics, clinical characteristics, neuropsychological scores and group differences for depressed bipolar patients and matched controls. *Significant at the 0.05 level. SD=standard deviation; df=degrees of freedom; HAM-17 Hamilton Depression Rating Scale 17-item; YMRS=Young Mania Rating Scale; DSST=Digit Symbol Substitution Test; SCOLP= Speed and Capacity of Language Processing; CPT=Continuous Performance Test; RAVLT=Rey Auditory Verbal Learning Test; SWM=Spatial Working Memory.*

	Healthy controls				Major Depressive Disorder				Group differences			
	N	**Mean [N]**	**SD [%]**	**Median**	**N**	**Mean [N]**	**SD [%]**	**Median**	**Statistic (*t*/*w*)**	***df***	***p***	**Cohen's *d***
Sex (N/% female)	41	[26]	[63]	-	41	[26]	[63]	-	χ^2=0.0	1	1	-
Age (years)	41	30.78	9.80	28.00	41	32.66	10.45	32.00	935.5	-	.380	0.19
Education (years)	40	15.31	2.63	15.50	41	13.63	2.84	13.00	534.5*	-	.006	0.62
HAM-17	-	-	-	-	41	20.61	4.20	21.00	-	-	-	-
National Adult Reading Test	41	109.88	6.89	111.00	41	107.98	10.94	110.00	805.5	-	.749	0.21
Processing Speed	**41**	**0.00**	**1.00**	**0.21**	**41**	**-0.24**	**0.92**	**-0.29**	**697.0**	**-**	**.184**	**0.26**
DSST (correct)	41	64.93	10.02	67.00	41	62.49	9.23	62.00	697.0	-	.184	0.26
Sustained Attention	**41**	**0.00**	**0.84**	**0.27**	**41**	**-1.59**	**2.47**	**-0.72**	**460.0***	**-**	**<.001**	**0.87**
Vigil CPT total errors	41	3.15	3.03	3.00	41	10.31	12.31	5.00	1145.0*	-	.005	0.81
Vigil CPT ex-Gaussian *tau*	30	67.71	30.99	65.92	39	90.34	58.01	79.95	735.0	-	.070	0.48
Executive Function	**41**	**0.00**	**0.97**	**0.02**	**41**	**0.13**	**0.90**	**0.12**	**0.63**	**79.6**	**.533**	**0.14**
ToL average perfect solutions	39	73.88	14.19	75.00	35	75.54	13.50	75.00	722.5	-	.666	0.12
ToL average excess moves	39	0.91	0.54	0.88	35	0.88	0.56	0.81	-0.19	70.6	.851	0.05
Verbal memory	**41**	**0.00**	**0.80**	**0.16**	**41**	**-0.38**	**0.74**	**-0.35**	**572.5***	**-**	**.013**	**0.50**
RAVLT A1 to 5 total	41	54.59	8.79	55.00	41	50.83	8.18	51.00	-2.00*	79.60	.049	0.45
RAVLT % retained at A6	41	0.89	0.18	0.92	41	0.85	0.15	0.86	691.0	-	.163	0.27
RAVLT % retained at A7	41	0.88	0.17	0.91	41	0.80	0.18	0.80	-2.06*	79.7	.043	0.46
Visuo-spatial memory	**41**	**0.00**	**0.72**	**0.10**	**41**	**-0.72**	**0.86**	**-0.69**	**-4.09***	**77.4**	**<.001**	**0.92**
Spatial Recognition % correct	41	84.76	10.84	85.00	41	78.23	11.34	80.00	568.0*	-	.011	0.60
Pattern Recognition % correct	41	90.29	9.57	91.67	41	82.42	14.27	83.33	563.5*	-	.010	0.66
SWM total errors	41	22.34	16.80	20.00	41	36.34	21.51	35.00	1150.0*	-	.004	0.73
SWM strategy count	41	32.37	5.79	34.00	41	35.90	6.13	36.00	2.686*	79.8	.009	0.60

*Table 7: Demographics, clinical characteristics, neuropsychological scores and group differences for Major Depressive Disorder patients and matched controls. *Significant at the 0.05 level. SD=standard deviation; df=degrees of freedom; HAM-17 Hamilton Depression Rating Scale 17-item; DSST=Digit Symbol Substitution Test; CPT=Continuous Performance Test; ToL=Towers of London; RAVLT=Rey Auditory Verbal Learning Test; SWM=Spatial Working Memory.*

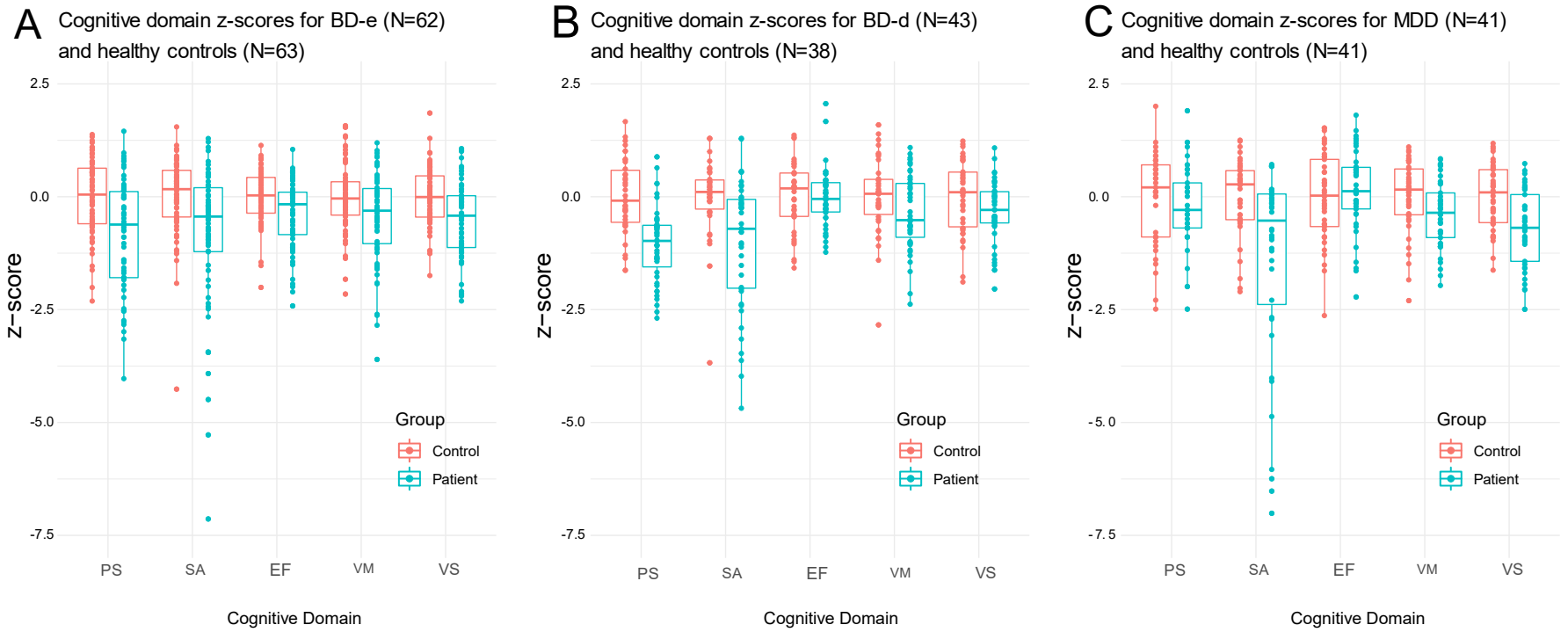

Figure 37: Cognitive domain z-scores for each dataset. A) Cognitive domain z-scores for BD-e and matched healthy controls. B) Cognitive domain z-scores for BD-d and matched healthy controls. C) Cognitive domain z-scores for MDD and matched healthy controls. BD-e=bipolar disorder euthymic; BD-d=bipolar disorder depressed; MDD=major depressive disorder; PS=processing speed; SA=sustained attention; EF=executive function; VM=verbal learning and memory; VS=visuo-spatial memory.

Group differences in each cognitive domain are illustrated in Figure 37. BD-e patients scored significantly worse than matched controls on most of the individual test scores except TMT-B, RAVLT A6 and A7, Pattern Recognition, and SWM strategy. BD-e patients had significantly worse performance than matched controls on all five of the domain scores. The ESs varied from medium to large, with PS showing a large ES (*d*=0.81), followed by moderate ESs for VS (*d*=0.69), SA (*d*=0.64), EF (*d*=0.57), and VM (*d*=.52). BD-d patients scored significantly worse than matched controls on most of the individual test scores except digit span backwards, RAVLT A6 and A7, Spatial Recognition, Pattern Recognition and SWM errors. BD-d patients had significantly worse performance than matched controls on PS and SA with large ESs (*d*=1.26 and *d*=0.85, respectively), but not on EF, VM or VS. MDD patients scored significantly worse than matched controls on most of the individual test scores except DSST, Vigil CPT ex-Gaussian *tau*, both ToL measures, and RAVLT A6. MDD patients scored significantly lower than matched controls on SA (*d*=0.87), VM (*d*=0.50), and VS (*d*=0.92), but not PS or EF.

4.3.2. Correlations between cognitive domains

Table 8 shows the results of Spearman's correlations between the cognitive domains for each dataset. Age and premorbid IQ were regressed out of the domain scores and the data were z-scored based on HC data, therefore scores represent deviations from the 'norm', with a lower z-score indicating poorer performance. For BD-e, PS, SA, EF, and VM were positively correlated with one another. None of the correlations for VS reached significance. For BD-d, PS and SA were positively correlated, and VS was positively correlated with SA, EF, and VM. For MDD, PS was positively correlated with SA and VS, and SA was positively correlated with VS.

The regularised partial correlation network graphs are shown in Figure 38 and illustrate relationships between the cognitive domains. For BD-e, most of the cognitive domains appeared to be interrelated. More relationships are present here than were significant in the Spearman's correlations, since we did not exclude network edges based on *p*-values. The strongest relationship was between SA and EF (*r*=.30), with moderate relationships also between EF and VM (*r*=.21), and SA and PS (*r*=.20). For BD-d, only seven edges survived regularisation, with the strongest relationships between VS and SA (*r*=.29), and VS and VM (*r*=.30). For MDD, only three relationships were present: a moderate relationship between SA and VS (*r*=.24), and the other two showing weak relationships (*r*≤.10).

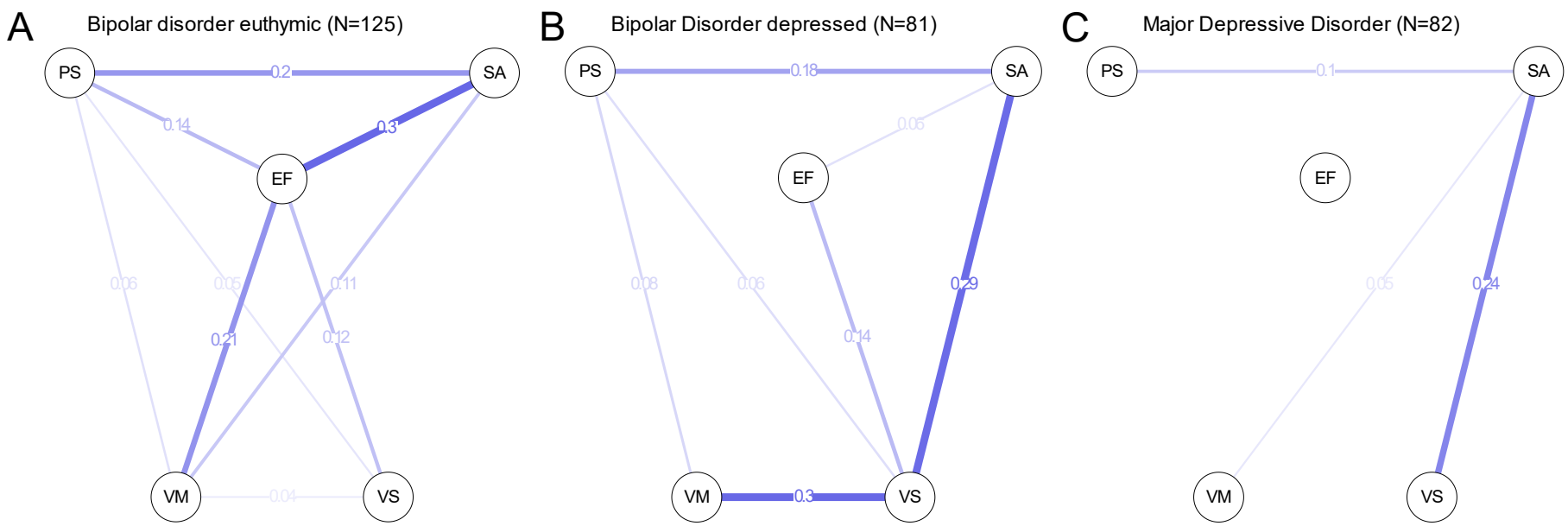

Figure 38: Networks of cognitive functioning for each dataset. Networks represent regularised partial correlations to estimate relationships between cognitive domains. Networks were estimated using LASSO penalty and EBIC model selection with hyperparameter gamma=0. The five nodes represent each cognitive domain and edges between them represent the relationship between two domains. Nodes were placed manually in a set layout, to allow for easier comparison of networks. Edges can be interpreted in the same way as partial correlations between each pair of cognitive domains. The strength of each relationship (i.e., the weight of the edge) is represented by the thickness of the line, where a thicker line indicated a stronger relationship. The partial correlation coefficient for each relationship is displayed on the edge between the two corresponding nodes. Each dataset contains patients and controls. A) Network for the bipolar disorder euthymic dataset. B) Network for the bipolar disorder depressed dataset. C) Network for the major depressive disorder dataset. PS=processing speed, SA=sustained attention, EF=executive function, VM=verbal learning and memory, VS=visuo-spatial memory.

		BD euthymic				BD depressed				MDD			
		SA	**EF**	**VM**	**VS**	**SA**	**EF**	**VM**	**VS**	**SA**	**EF**	**VM**	**VS**
PS	*df*	123	123	123	123	79	79	79	79	80	80	80	80
	rho	.306*	.303*	.192*	.155	.279*	.055	.175	.184	.311*	-.037	.031	.223*
	p	.001	.001	.032	.083	.012	.628	.117	.101	.004	.740	.779	.044
SA	*df*		123	123	123		79	79	79		80	80	80
	rho		.388*	.282*	.004		.137	.214	.446*		.022	.214	.430*
	p		<.001	.001	.968		.223	.055	<.001		.842	.053	<.001
EF	*df*			123	123			79	79			80	80
	rho			.327*	.170			.143	.224*			-.125	.052
	p			<.001	.057			.202	.045			.262	.643
VM	*df*				123				79				80
	rho				.101				.403*				.182
	p				.261				<.001				.102

*Table 8: Results of Spearman's correlations between each pair of cognitive domains for each dataset. *Significant at the .05 level. BD=bipolar disorder; MDD=major depressive disorder; PS=processing speed; SA=sustained attention; EF=executive function; VM=verbal learning and memory; VS=visuo-spatial memory; df=degrees of freedom.*

4.3.3. Mediation models

Group differences in the cognitive domains were tested in section 4.3.1. For completeness, and to allow comparison of statistics to test for mediation, simple regressions were run to evaluate relationships between group status, core cognitive variables, and memory. Note that these results may differ slightly from the group differences reported in in section 4.3.1 since non-parametric tests were used in some cases and group differences were conducted before imputing missing data and accounting for age and premorbid IQ. Multiple regressions were run to test whether the effect of group on memory remained significant after accounting for each core cognitive domain. Separate models were run with either PS, SA, or EF as the mediator and either VM or VS as the dependent variable. Results of simple linear regressions to test the effect of each core cognitive domain on group status for each dataset are presented in Table 9. Results of simple and multiple regressions with either VM or VS as the outcome variable are presented in Table 10 and Table 11, respectively.

For BD-e, group significantly predicted VM and PS, and PS predicted VM. However, PS did not mediate the effect of group on VM, since PS did not have a significant effect in this model. Group significantly predicted SA, and SA predicted VM. Multiple regression suggested that SA partially mediated the relationship between group and VM, which was confirmed with the Sobel test (S=-2.01, SE=659.32, *p*=.044). Group significantly predicted EF, and EF predicted VM. EF partially mediated the relationship between group and VM (Sobel test: S=-2.43,

SE=1.20, *p*=.015). Group significantly predicted VS, PS, and SA, however, neither PS nor SA predicted VS. Group significantly predicted EF, and EF predicted VS. EF partially mediated the relationship between group and VS (Sobel test: S=-2.54, SE=0.02, *p*=.011). Therefore, SA and EF, separately, partially mediated the effect of group status on VM, and EF partially mediated the effect of group status on VS.

For the BD-d dataset, group did not have a significant effect on VM, or VS, so mediation analyses were not interpreted. For the MDD dataset, group had a significant effect on both VM and VS, but only SA had a significant effect on group status. SA significantly predicted VM and VS. The multiple regression to test the effects of group and SA on VM was significant overall, however main effects of group and SA were not significant. SA partially mediated the relationship between group and VS (Sobel test: S=2.41, SE=0.10, *p*=.016). There was no clear evidence that PS or EF individually mediated the effect of group on VM or VS for either BD-d or MDD datasets.

Dataset	**Model**	***F***	***df***	***p***	R^2	***β***	***t***
Bipolar Disorder euthymic	PS	24.21*	1, 123	<.001	0.17	-0.41	-4.92
	SA	15.35*	1, 123	<.001	0.11	-0.33	-3.92
	EF	12.94*	1, 123	<.001	0.10	-0.31	-3.60
Bipolar Disorder depressed	PS	33.23*	1, 79	<.001	0.30	-0.54	-5.77
	SA	14.67*	1, 79	<.001	0.16	-0.40	-3.83
	EF	0.00	1, 79	.946	0.00	-0.01	-0.07
Major Depressive Disorder	PS	0.43	1, 80	.515	0.01	-0.07	-0.65
	SA	20.54*	1, 80	<.001	0.20	-0.45	-4.53
	EF	0.59	1, 80	0.443	0.01	0.09	0.77

*Table 9: Results of simple linear regressions to test the effect of each core cognitive domain (PS, SA, and EF) on group status (patient vs control), for each dataset. *Significant at the .05 level. df=degrees of freedom; PS=processing speed, SA=sustained attention, EF=executive function.*

	Verbal learning and memory								
	Model	***F***	***df***	***p***	***R²***	***R²* adj.**	***IV***	***β***	***t***	***p***
BD euthymic	Group	10.51*	1, 123	.002	.08	.07		-0.28*	-3.24	
	PS	5.79*	1, 123	.018	.04	.04		0.21*	2.41	
	SA	10.30*	1, 123	.002	.08	.07		0.28*	3.21	
	EF	17.08*	1, 123	.000	.12	.11		0.35*	4.13	
	Group + PS	6.06*	2, 122	.003	.09	.08	Group	-0.23*	-2.47	.015
							PS	0.12	1.24	.216
	Group + SA	8.09*	2, 122	.001	.12	.10	Group	-0.21*	-2.34	.021
							SA	0.21*	2.30	.023
	Group + EF	11.12*	2, 122	<.001	.15	.14	Group	-0.19*	-2.18	.034
							EF	0.29*	3.32	.001
	Model	***F***	***df***	***p***	***R²***	***R²* adj.**	***IV***	***β***	***t***	***p***
BD depressed	Group	2.62	1, 79	.110	.03	.02		-0.18	-1.62	
	PS	3.06	1, 79	.084	.04	.03		0.19	1.75	
	SA	2.83	1, 79	.097	.03	.02		0.19	1.68	
	EF	1.19	1, 79	.278	.01	.00		0.12	1.09	
	Group + PS	1.84	2, 78	.166	.05	.02	Group	-0.10	-0.80	.429
							PS	0.14	1.03	.305
	Group + SA	1.95	2, 78	.149	.05	.02	Group	-0.13	-1.04	.302
							SA	0.14	1.13	.261
	Group + EF	1.91	2, 78	.156	.05	.02	Group	-0.18	-1.61	.111
							EF	0.12	1.09	.279
	Model	***F***	***df***	***p***	***R²***	***R²* adj.**	***IV***	***B***	***t***	***p***
MDD	Group	5.17*	1, 80	.026	.06	.05		-0.25*	-2.27	
	PS	0.00	1, 80	.947	.00	-.01		0.01	0.07	
	SA	4.53*	1, 80	.036	.05	.04		0.23*	2.13	
	EF	1.40	1, 80	.241	.02	.00		-0.13	-1.18	
	Group + PS	2.56	2, 79	.084	.06	.04	Group	-0.25*	-2.26	.027
							PS	-0.01	-0.10	.923
	Group + SA	3.38*	2, 79	.039	.08	.06	Group	-0.18	-1.47	.145
							SA	0.15	1.25	.216
	Group + EF	3.11	2, 79	.050	.07	.05	Group	-0.24*	-2.18	.032
							EF	-0.11	-1.02	.312

*Table 10: Results of simple and multiple regression models to test whether the core cognitive domains mediate the effect of Group on verbal learning and memory. For each dataset, simple regressions are presented in the first four rows to test the effect of Group and each core cognitive variable individually on verbal learning and memory. The last three rows represent multiple regressions to test whether each core cognitive variable can account for the group difference in verbal learning and memory. *Significant at the .05 level. df=degrees of freedom; adj.=adjusted; IV=independent variable; BD=bipolar disorder; MDD=major depressive disorder; PS=processing speed, SA=sustained attention, EF=executive function, VM=verbal learning and memory, VS=visuo-spatial memory.*

	Visuo-spatial memory									
	Model	***F***	***df***	***p***	***R^2***	***R^2* adj.**	***IV***	***β***	***t***	***p***
BD euthymic	Group	6.11*	1, 123	.015	.05	.04		0.22*	2.47	
	PS	2.86	1, 123	.093	.02	.01		0.15	1.69	
	SA	0.06	1, 123	.805	.00	-.01		0.02	0.25	
	EF	6.25*	1, 123	.014	.05	.04		0.22*	2.50	
	Group + PS	7.98*	2, 123	.001	.12	.10	Group	0.33*	3.58	<.001
							PS	0.29*	3.07	.003
	Group + SA	3.72*	2, 123	.027	.06	.04	Group	0.25*	2.71	.008
							SA	0.11	1.14	.225
	Group + EF	9.79*	2, 123	<.001	.14	.12	Group	0.32*	3.57	.001
							EF	0.32*	3.59	<.001
	Model	***F***	***df***	***p***	***R^2***	***R^2* adj.**	***IV***	***β***	***t***	***p***
BD depressed	Group	1.33	1, 79	.252	.02	.00		-0.13	-1.15	
	PS	4.29*	1, 79	.042	.05	.04		0.23*	2.07	
	SA	15.66*	1, 79	<.001	.17	.15		0.41*	3.96	
	EF	4.77*	1, 79	.032	.06	.04		0.24*	2.18	
	Group + PS	2.12	2, 78	.127	.05	.03	Group	-0.01	-0.06	.955
							PS	0.22	1.70	.094
	Group + SA	7.80*	2, 78	.001	.17	.15	Group	0.04	0.34	.735
							SA	0.42*	3.75	<.001
	Group + EF	3.07	2, 78	.052	.07	.05	Group	-0.13	-1.16	.248
							EF	0.24*	2.18	.032
	Model	***F***	***df***	***p***	***R^2***	***R^2* adj.**	***IV***	***β***	***t***	***P***
MDD	Group	13.51*	1, 80	<.001	.14	.13		-0.38*	-3.68	
	PS	3.17	1, 80	.079	.04	.03		0.20	1.78	
	SA	17.52*	1, 80	<.001	.18	.17		0.42*	4.19	
	EF	0.05	1, 80	.816	.00	-.01		0.03	0.23	
	Group + PS	8.25*	2, 79	.001	.17	.15	Group	-0.37*	-3.59	.001
							PS	0.17	1.64	.105
	Group + SA	11.42*	2, 79	<.001	.22	.20	Group	-0.24*	-2.13	.036
							SA	0.32*	2.85	.006
	Group + EF	6.86*	2, 79	.002	.15	.13	Group	-0.39*	-3.70	<.001
							EF	0.06	0.57	.572

*Table 11: Results of simple and multiple regression models to test whether the core cognitive domains mediate the effect of Group on visuo-spatial memory. For each dataset, simple regressions are presented in the first four rows to test the effect of Group and each core cognitive variable individually on visuo-spatial memory. The last three rows represent multiple regressions to test whether each core cognitive variable can account for the group difference in visuo-spatial memory. *Significant at the .05 level. df=degrees of freedom; adj.=adjusted; IV=independent variable; BD=bipolar disorder; MDD=major depressive disorder; PS=processing speed, SA=sustained attention, EF=executive function, VM=verbal learning and memory, VS=visuo-spatial memory.*

4.3.4. Hierarchical regressions

Thus far, we have established whether core cognitive functions play a role in the relationship between group status and memory separately, but not when the core cognitive functions are considered together. Hierarchical regression models were run to explore the effect of core cognitive functioning on the group difference in memory when they were considered together. Results are summarised for BD-e, BD-d, and MDD datasets in Table 12, Table 13, and Table 14, respectively. The full results tables can be found in Appendix B (Table 26, Table 27, Table 28, Table 29, Table 30, and Table 31).

For BD-e, the full model of the effect of PS, SA, EF, and group status on VM was statistically significant overall and explained 14% of the variance in VM, $F(4,120)=5.97$, $p<.001$, adjusted $R^2=.14$. The *F*-change results of the hierarchical regressions showed that, once EF was accounted for, PS and SA did not add anything of statistical significance to the models. This was only apparent when adding EF at the first step; when PS or SA were entered first, they appeared to contribute to the model. However, it was only after both EF and PS were accounted for that group status no longer had a significant effect on VM, suggesting that EF and PS together may explain the between-group difference in VM.

The full model of the effect of PS, SA, EF, and group status on VS was statistically significant overall and explained 15% of the variance in VS, $F(4,120)=6.66$, $p<.001$, adjusted $R^2=.15$. When SA was entered first, it appeared to contribute to the model, however, once EF and PS were accounted for, SA did not add anything of statistical significance. Adding PS and EF had a significant effect on the model. After accounting for all three core cognitive variables, the effect of group remained significant, suggesting that PS and EF do not completely account for the between-group difference in VS. A post-hoc test was run to assess the R^2-change between two models: the effect of group on VM, and the effect of PS, EF, and group on VM. The *F*-change statistic was significant (*F*-change=2.69, $p<.001$), and the model with PS, EF, and group explained 13% more of the variance in VS than group alone.

There was no group difference in VM or VS for the BD-d dataset. The full model of the effect of PS, SA, EF, and group status on VM not significant overall, $F(4,76)=1.39$, $p=.247$; none of the predictors appeared to influence VM. The full model of the effect of PS, SA, EF, and group status on VS was significant overall, $F(4,76)=5.27$, $p=.001$, adjusted $R^2=.22$. This appeared to

be driven by the relationship between SA and VM; SA was the only variable that had a significant effect on the model.

For the MDD dataset, the full model of the effect of PS, SA, EF, and group status on VM was not significant, $F(4,77)=2.06$, $p=.094$. The only significant model was when SA and group status predicted VM, $F(4,76)=3.38$, $p=.039$, adjusted $R^2=.06$, however, neither SA nor group status had a significant effect on VM individually. The full model of the effect of PS, SA, EF, and group status on VS was significant: PS, SA EF and group status explained 20% of the variance in VS, $F(4,77)=5.92$, $p<.001$, adjusted $R^2=.20$. The hierarchical regression showed that, once SA was accounted for, PS and EF did not add anything of statistical significance to the models. This was only apparent when adding SA at the first step; when PS or EF were entered first, they appeared to contribute to the model. After accounting for all three core cognitive variables, or SA alone, the effect of group remained significant, suggesting that SA does not completely account for the between-group difference in VS.

Bipolar Disorder euthymic		Verbal learning and memory				Visuo-spatial memory			
		Model	Change Statistics			Model	Change Statistics		
Model		R^2 *adj.*	R^2 change	*F* change	*p*	R^2 *adj.*	R^2 change	*F* change	*p*
1	PS + Group	.08	.09	6.06*	.003	.10	.12	7.98*	.001
	PS + SA + Group	.10	.03	-0.51*	.042	.10	<.01	-2.63	.652
	PS + SA + EF + Group	.14	.05	0.42*	.012	.15	.06	1.30*	.003
2	SA + Group	.10	.12	8.09*	.001	.04	.06	3.72*	.027
	SA + PS + Group	.10	<.01	-2.54*	.042	.10	.06	1.64*	.005
	SA + PS + EF + Group	.14	.05	0.42*	.012	.15	.06	1.30*	.003
3	EF + Group	.14	.15	11.18*	<.001	.12	.14	9.79*	<.001
	EF + PS + Group	.14	<.01	-3.64	.552	.16	.04	-0.99*	.016
	EF + PS + SA + Group	.14	.01	-1.57	.271	.15	<.01	-2.14	.544

*Table 12: Results of hierarchical regression analyses examining the ability of core cognitive functions to account for between-groups variance in verbal learning, and memory and visuo-spatial memory performance for the bipolar disorder euthymic dataset. Three hierarchical models are presented, each with three steps. In model 1, PS and Group were entered first, followed by the addition of SA, then EF. In model 2, SA and Group were entered first, followed by addition of PS, then EF. For model 3, EF and Group were entered first, followed by addition of PS, then SA. Note that the R^2 change statistic for the first line in each model is just R^2. *Significant at the 0.05 level. PS=processing speed; SA=sustained attention; EF=executive function.*

Bipolar Disorder depressed		Verbal learning and memory				Visuo-spatial memory			
		Model	Change Statistics			Model	Change Statistics		
Model		R^2 adj.	R^2 change	*F* change	*p*	R^2 adj.	R^2 change	*F* change	*p*
1	PS + Group	.02	.05	1.84	.166	.03	.05	2.12	.127
	PS + SA + Group	.02	.01	-0.25	.302	.16	.14	3.88*	.001
	PS + SA + EF + Group	.02	.01	-0.20	.375	.18	.03	-0.73	.105
2	SA + Group	.02	.05	1.95	.149	.15	.17	7.80*	.001
	SA + PS + Group	.02	.01	-0.37	.356	.16	.02	-1.81	.146
	SA + PS + EF + Group	.02	.01	-0.20	.375	.18	.03	-0.73	.105
3	EF + Group	.02	.05	1.91	.156	.05	.07	3.07	.052
	EF + PS + Group	.02	.01	-0.29	.312	.07	.03	-0.01	.092
	EF + PS + SA + Group	.02	.01	-0.23	.400	.18	.11	2.21*	.002

*Table 13: Results of hierarchical regression analyses examining the ability of core cognitive functions to account for between-groups variance in verbal learning, and memory and visuo-spatial memory performance for the bipolar disorder depressed dataset. Three hierarchical models are presented, each with three steps. In model 1, PS and Group were entered first, followed by the addition of SA, then EF. In model 2, SA and Group were entered first, followed by addition of PS, then EF. For model 3, EF and Group were entered first, followed by addition of PS, then SA. Note that the R^2 change statistic for the first line in each model is just R^2. *Significant at the 0.05 level.PS=processing speed; SA=sustained attention; EF=executive function.*

Major Depressive Disorder		Verbal learning and memory				Visuo-spatial memory			
		Model	Change Statistics			Model	Change Statistics		
Model		R^2 *adj.*	R^2 change	*F* change	*p*	R^2 *adj.*	R^2 change	*F* change	*p*
1	PS + Group	.04	.06	2.56	.084	.15	.17	8.25*	.001
	PS + SA + Group	.05	.02	-0.25	.189	.20	.06	-0.37*	.016
	PS + SA + EF + Group	.05	.02	-0.25	.259	.20	<.01	-1.96	.612
2	SA + Group	.06	.08	3.38*	.039	.20	.22	11.42*	<.001
	SA + PS + Group	.05	<.01	-1.08	.636	.20	.01	-3.54	.365
	SA + PS + EF + Group	.05	.02	-0.25	.259	.20	<.01	-1.96	.612
3	EF + Group	.05	.07	3.11	.050	.13	.15	6.86*	.002
	EF + PS + Group	.04	<.01	-1.05	.870	.15	.03	-1.24	.097
	EF + PS + SA + Group	.05	.02	0.01	.161	.20	.06	0.30*	.019

*Table 14: Results of hierarchical regression analyses examining the ability of core cognitive functions to account for between-groups variance in verbal learning, and memory and visuo-spatial memory performance for the major depressive disorder dataset. Three hierarchical models are presented, each with three steps. In model 1, PS and Group were entered first, followed by the addition of SA, then EF. In model 2, SA and Group were entered first, followed by addition of PS, then EF. For model 3, EF and Group were entered first, followed by addition of PS, then SA. Note that the R^2 change statistic for the first line in each model is just R^2. *Significant at the 0.05 level. PS=processing speed; SA=sustained attention; EF=executive function.*

4.3.5. Exploratory analysis of patient and control groups separately

Thus far, we have analysed each dataset as a whole by grouping together patients and controls and using group status as a variable. However, it may be the case that patients and controls have different cognitive profiles with different relationships between cognitive functions. As an exploratory step, we used network graphs to assess relationships between cognitive domains in patients and controls separately, in each dataset. The networks are presented in Figure 39 and were constructed in the same way as described previously (section 4.2.4.3).

For BD-e patients, the network included associations between most cognitive domains and was similar to the network for the whole sample. PS was related to VS, EF was related to VM and VS, and SA was related to EF, which was in line with the results from the hierarchical regressions, suggesting that PS and EF may have effects on memory in BD-e patients. However, the control group showed an empty network where none of the relationships between cognitive domains survived regularisation. Conversely, for BD-d and MDD datasets, the patient groups had empty networks, whereas the control groups showed some associations between cognitive domains. For the BD-d control group, strong positive relationships appeared between VS and VM (r=.50), and VS and SA (r=.60). Moderate positive relationships were found between VM and PS (r=.28), VM and EF (r=.22), and EF and SA (r=.26). There appeared to be a negative relationship between VM and SA (r=-.33). For MDD, there was a positive relationship between PS and SA (r=.43), and a moderate relationship between SA and VS (r=.20).

We also ran multiple regression to test whether PS, SA, and EF predicted memory after accounting for the other core cognitive variables in each group. Results are presented in Table 32 in Appendix B and are generally in line with what was found in the network graphs: for BD-e patients, EF appeared to predict both VM and VS, and SA predicted VS. However, for the BD-e control group, the models were not significant, suggesting no effect of PS, SA, or EF on VM or VS. For BD-d patients and MDD patients the models were not significant. For the BD-d control group, SA appeared to predict VS. All other models were not statistically significant.

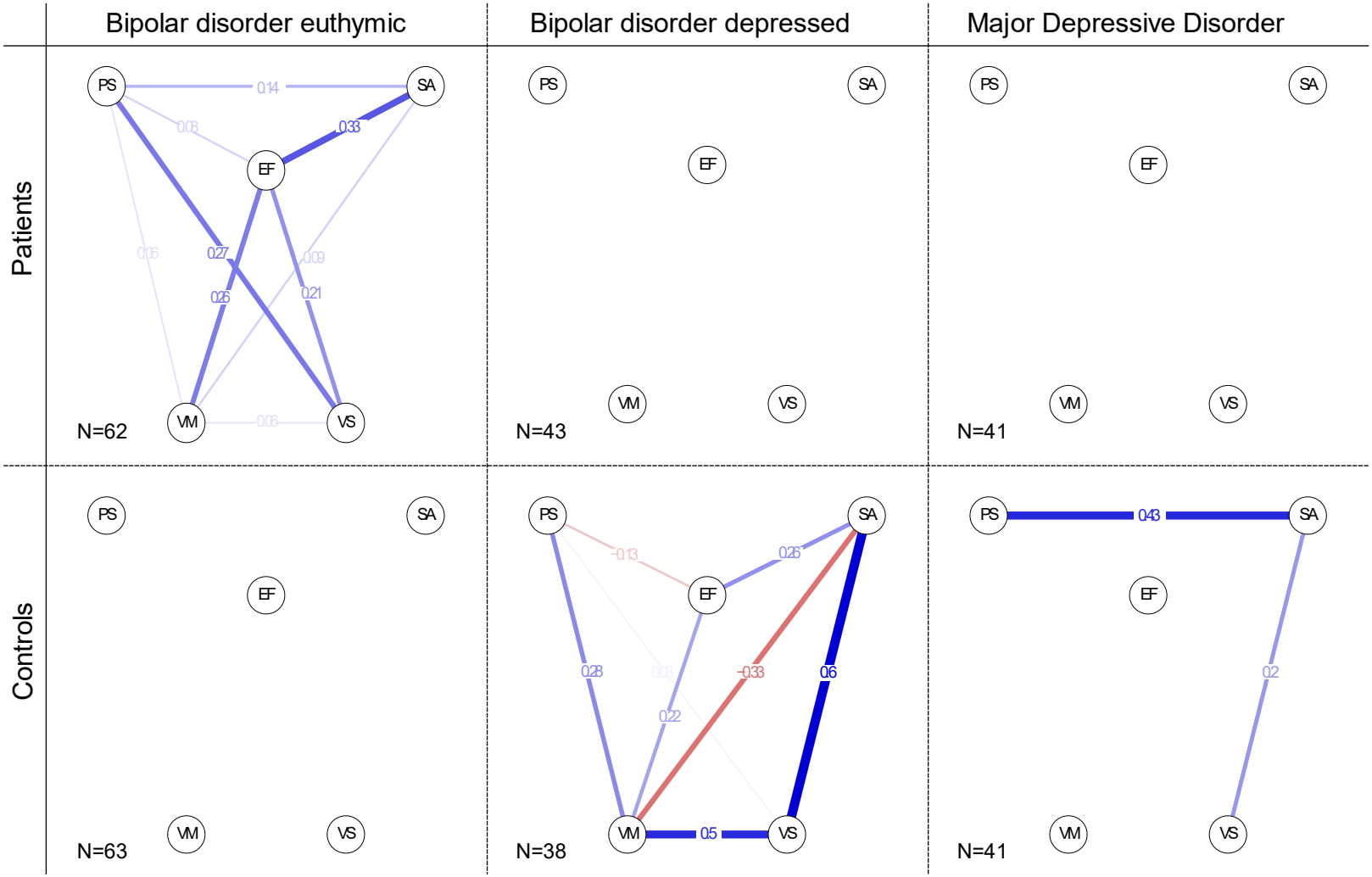

Figure 39: Networks of cognitive functioning for patient and control groups in each dataset. Networks represent regularised partial correlations estimated using LASSO penalty and EBIC model selection with hyperparameter gamma=0. The five nodes represent each cognitive domain and edges between them represent the relationship between two domains. Nodes were placed manually in a set layout, to allow for easier comparison of networks. The strength of each relationship is represented by the thickness of the line, where a thicker line indicated a stronger relationship. The partial correlation coefficient for each relationship is displayed on the edge between the two corresponding nodes. PS=processing speed, SA=sustained attention, EF=executive function, VM=verbal learning and memory, VS=visuo-spatial memory.

4.4. DISCUSSION

4.4.1. Summary of results

The results of the hierarchical regression models suggested that in euthymic BD patients, EF and, to some extent, PS explained the impairment in VM and VS. The depressed BD sample did not appear to differ from HCs on either VM or VS, so we could not assess whether PS, SA, or EF account for impairments in memory. In MDD patients, SA appeared to contribute to the impairment in VS memory, however, it did not completely account for the group difference. We found different results when the core cognitive functions were considered together compared to when they were considered as mediators alone, and the order of entry into the hierarchical regression mattered. For example, SA partially mediated the effect of group on VM in BD-e dataset on its own, but when PS and EF were accounted for, SA did not significantly contribute to the model. This highlights the importance of accounting for relationships between cognitive functions when researching cognitive impairment in mood disorders.

The networks for each group suggested that the relationships between cognitive functions may be different in BD-e, BD-d, and MDD patients, however, the exact cognitive profile of each group is not yet clear. We performed exploratory analysis on patient and control groups separately to see if cognitive profiles differed between patients and controls. Empty networks were produced for BD-d patients and MDD patients, whereas cognitive functions appeared to be interrelated in BD-e patients. Notably, each control group showed a different profile of cognitive interrelations, with one of the three HC groups showing an empty network. This may have resulted from methodological differences in each study, including the neuropsychological tests used to calculate each cognitive domain score, or perhaps differences in the control group in terms of demographic characteristics (although this was not formally tested). Alternatively, it could indicate that relationships between cognitive functions are unstable in control groups, or that this network analysis method is not reliable for this type of data or for our sample sizes.

4.4.2. Bipolar disorder euthymia

Our results concerning BD-e extend those from Thompson et al. (2009). While we used the same sample as this study, our methods differed in terms of the neuropsychological scores used to create the cognitive domain scores and the inclusion of SA as a potential core function.

However, our results indicated that SA may not play a role in memory in euthymic BD and confirm what was previously found in this dataset; PS and EF appeared to explain the group difference in memory. This is in line with other research that found that motor speed predicted deficits in wider cognitive functioning in BD (Salazar-Fraile et al., 2009). In our BD-e sample, PS showed the largest ES in the group differences, consistent with other studies reporting that PS had the largest impairment in BD (Bo et al., 2017). Where previous research failed to find a differential impairment in EF in euthymic BD, authors consequently concluded that there exists a generalised cognitive impairment in BD (Mann-Wrobel et al., 2011). Our study likewise found that the impairment in EF was not the largest ES in terms of cognitive impairments, however our following analysis demonstrated the role of EF in memory impairments in these patients, demonstrating the importance of considering the relationships between cognitive functions. In the network for BD-e patients, relationships existed between cognitive domains, whereas their matched control group had an empty network. This suggests more overlap of cognitive functions in euthymic BD than in a healthy model, which is in line with previous research in BD (Gallagher et al., 2014; Gallagher, Gray, et al., 2015).

4.4.3. Bipolar disorder depression

We did not find an impairment in VM or VS for the depressed BD sample, despite moderate ESs. Our sample size was modest, so it is possible that this is a type-II error and a significant difference may be found in a larger sample. Indeed, significant differences in tests of VM and VS were found in an earlier analysis of this dataset with a larger sample (Gallagher et al., 2014). The analysis by Gallagher et al. demonstrated that the proportion of patients who were impaired on these tests compared to controls changed depending on the cut-off score: for example, only 2% of patients showed impairments on Spatial Recognition at the 5^{th} percentile cut-off, compared to 23% of patients showing impairments at the 10^{th} percentile cut-off. Results may therefore be driven by inter-individual variability in performance across measures. The null finding may also be due to the neuropsychological tests that were used to create domain scores. We used the percent retained from the RAVLT A6 and A7 trials to measure verbal memory, to attempt to remove the learning component of the score (i.e., the number of words initially encoded). Previous research similarly failed to find a difference in VM or VS between MDD and controls using the same scoring method (Nilsson et al., 2016). The null findings are also in line with Salazar-Fraile et al. (2009) who did not find evidence that motor speed predicted VM.

Interestingly, networks and regression analysis did not detect overlaps between cognitive functions in our depressed BD patients, despite the fact that this dataset was taken from the same sample as was used by Gallagher et al. (2014), who did find evidence for overlap. We used a subset of this sample, including only participants with sufficient cognitive data, resulting in a smaller sample than the one reported previously, which may explain this discrepancy. Further, we used a different selection of cognitive tests to match the domain scores as closely as possible to the other datasets in our analysis, including the use of ex-Gaussian parameters that were not used previously. Gallagher et al. (2014) used PCA, which may be more sensitive to latent cognitive variables than regression analysis. Previous analysis using network graphs to illustrate the cognitive profile of depressed BD patients found that the BD network was less densely inter-connected that the MDD network and that EF was central in the BD network (Galimberti et al., 2020). We found that EF was central to the cognitive profile in euthymic BD patients, but not depressed BD patients. In our BD-d sample, the strongest connections appeared between VS and VM, and VS and SA. The differences in these results may be due to several factors: Galimberti et al. (2020) did not include a HC group, whereas our cognitive domain scores represent a deviation from the average control score. The analysis by Galimberti et al. did not include PS or SA; the addition of these domains may have altered the networks.

4.4.4. Major depressive disorder

Our results are consistent with previous research showing an impairments in SA in MDD (Schmidt et al., 2021). Our study further suggested that SA may play a role in memory impairments in MDD. This was the first study, to our knowledge, that tested the role of SA in wider cognitive dysfunction in mood disorder groups. Nilsson et al. (2016) found that EF and attention characterised cognitive impairments in MDD but found that the attentional deficit persisted after controlling for EF, but not vice versa. This is similar to our result, suggesting that attention may be a primary deficit in MDD that may lead to secondary impairments. However, the tasks used to measure attention in Nilsson et al. (2016) may be considered measures of PS, so these results require replication.

In MDD, once SA was accounted for, PS and EF did not add anything of significance to explain the impairment in memory in MDD. There is limited research on the effect of PS on cognition in MDD. One study likewise failed to find evidence that PS can explain general cognitive

impairment in MDD (McDermott & Ebmeier, 2009). PS and EF has been consistently found to mediate the depression-related cognitive dysfunction in LLD (Butters et al., 2004; Sexton et al., 2012; Sheline et al., 2006); our results perhaps suggest that these functions may not play a role in younger adults. However, our results are in contrast with other research showing that PS and EF may be primarily impaired in MDD and may cause secondary cognitive impairments (J. Liu et al., 2019; Zaremba et al., 2019). Previous studies also found that PS, EF, attention and memory were highly correlated with one another in MDD and HCs (J. Liu et al., 2019). This contrasts with our study, where PS only correlated with SA and VS, and SA correlated with VS; EF and VM were not correlated with any other domain score. Previous network analysis of a MDD sample suggested that memory was central in the MDD network (Galimberti et al., 2020), whereas our results only showed a moderate connection between VS and SA, and a strong connection between PS and SA. However, differences in methodology mentioned above should be considered when comparing these results. Importantly, no previous study included SA in their analysis, so our findings perhaps suggest that SA may supersede the role of PS and EF in wide cognitive impairment in MDD. This is the first study to our knowledge to assess the role of SA as a core cognitive function in MDD and suggests that more research is required to fully explore the role of SA on wider functioning in MDD.

4.4.5. Limitations

There are some important limitations to our results. Firstly, datasets were used from previous studies and thus were not initially designed to answer our research questions. The datasets are cross-sectional, so causation cannot be inferred from the correlational results. Due to the exploratory nature of our analysis, we did not control for multiple comparisons, so our results may be susceptible to type-I errors and should be replicated. We also had modest sample sizes, with a larger BD-e sample than BD-d and MDD. Many of the relationships reported did not reach significance at the 0.05 level but were approaching significance; this may indicate an issue with statistical power. Future research with larger samples may benefit from employing multivariate and dimension reduction techniques to mitigate these issues.

Missing data were not consistent across datasets: for MDD, a larger proportion of data was missing for ToL and Vigil *tau*. Missing data were imputed based on demographic and neuropsychological data and imputed scores were not considered extreme or unusual. The individual neuropsychological scores were combined into composite scores to represent

domains, so imputed data were combined with real scores, which may mitigate the risks of imputing data to a degree. However, imputing missing data with regression using the other cognitive variables may have artificially inflated relationships between cognitive variables (Raymond & Roberts, 1987). Therefore, results concerning SA and EF in the MDD dataset should be interpreted with caution and replication of our results is required.

4.4.5.1. Neuropsychological methodology

There are some limitations concerning the neuropsychological assessment in our study, which differed between the datasets. Firstly, we created composite scores to represent cognitive domains in a theory-driven way, rather than using data-driven techniques such as PCA to identify latent components. This was done to improve interpretability of the results, to allow for comparison to previous research (Nilsson et al., 2016; Thompson et al., 2009), and because our samples were small, which can pose problems for dimension-reduction techniques. Further, the variety of cognitive tests used in each dataset may have resulted in different principal components being found in each dataset, which would have complicated interpretation and limited comparison between patient groups. However, creating composite scores from manually chosen tests leaves the scores vulnerable to the task impurity problem, where a test thought to measure a distinct cognitive function likely also captures elements of other cognitive functions. We attempted to mitigate this where possible, by removing elements of PS and attention from EF tests (e.g., for TMT-B, Stroop, and digit span). However the problem may persist for other measures: for example, DSST represented PS, but it likely captures elements of attention and WM (Joy et al., 2004). SWM was grouped with tests of VS here, but it also captures elements of EF (P. J. Smith et al., 2013). These confounds may have led to potentially exaggerated relationships between cognitive domains, particularly between EF and VS domain scores. This adds a further challenge in interpreting the results and assessing the hierarchical nature of cognitive functions.

Secondly, PS, EF, and VS domains were not comparable between datasets. PS was somewhat similar: all three datasets included the DSST, and we included other measures of PS where possible. The datasets also shared most tests of VS memory, with some minor differences. However, the EF domain differed across datasets. EF can be split into three sub-components: shifting of mental set, monitoring and updating of WM, and inhibiting of prepotent responses (Miyake, Friedman, et al., 2000). We aimed to include tests of each sub-component, but this

was only possible for BD-e. The MDD dataset did not share any test of EF with the two BD datasets and instead used the ToL task, which is thought to tests planning and may not capture all three subcomponents of EF (Humes et al., 1997). However, previous work has shown that mood disorder patients do not show a differential deficit on any of the three subcomponents, so executive impairments should be detected by any EF test (Paelecke-Habermann et al., 2005; Thompson et al., 2009). Veral fluency was used to measure EF in the BD-d dataset, however, while some research uses verbal fluency tasks to measure EF (Thompson et al., 2005; Watson et al., 2012), fluency is often considered to measure PS (Henry & Crawford, 2005). It may therefore be the case that the use of different tests of EF may have confounded results.

4.4.5.2. Demographic and clinical confounds

There are also demographic and clinical confounds that may have affected our results. Our samples are difficult to compare due to important differences in methodological and clinical factors: the MDD group were unmedicated and we did not have information about any comorbidities present for BD-d. We tested samples of euthymic and depressed BD patients, but we did not formally compare the groups, so differences discussed are only descriptive. We were not able to investigate the cognitive profile of MDD in remission. Therefore, while we were able to analyse data in different mood states, future research is necessary to formally compare differences between them.

We controlled for age and premorbid IQ in our analysis, as these are known to confound cognitive performance (Hu et al., 2022; Machado et al., 2018; Tsitsipa & Fountoulakis, 2015). We did not control for education here, as education is likely to overlap with premorbid IQ and may be influenced by clinical factors such as age of illness onset in patients (Fergusson & Woodward, 2002; Sørensen et al., 2012). However, we did not control for other clinical characteristics that may influence cognitive performance. For example, increased illness severity and current or previous psychosis have been linked to cognitive impairment in mood disorders (Bora et al., 2007; Gorwood et al., 2008). There may also be differential effects of specific clinical characteristics on different cognitive functions: Nilsson et al. (2016) found a relationship between attention illness duration, but not severity, on attention in MDD. Other research found that cognitive impairments are not explained by history of psychosis or number of previous episodes, suggesting that clinical characteristics may not completely confound the results (Demmo et al., 2016). Childhood trauma is also related to cognitive

impairment in mood disorders (Bücker et al., 2013; Jørgensen et al., 2023), therefore the presence of childhood trauma in each sample could have affected the cognitive scores. Future research should therefore assess the effect of clinical factors such as age of onset, illness duration, presence of psychosis, and childhood trauma, on the cognitive profile of BD and MDD.

4.4.6. Conclusion and interim summary

We used hierarchical regression to investigate the role of core cognitive functions on memory in BD and MDD. Overall, our results suggest that PS, SA, and EF may explain some of the memory impairments found in mood disorders. SA may be more central in depressed states. While our results require replication to assess the nature of the interrelationships between cognitive functions in mood disorders (and in healthy controls), our study justifies further research into the concept of a hierarchy of cognitive dysfunction in BD and MDD by highlighting potentially important relationships between cognitive domains. Future research should examine the role of core cognitive functions in BD and MDD in different mood states, controlling for potential clinical confounds. It would also be of interest to examine the effects of core cognitive functions on wider functioning other than memory, for example, language or other visuospatial abilities.

Chapter 5 NEURAL CORRELATES OF CORE COGNITIVE IMPAIRMENT IN BIPOLAR DISORDER

5.1. INTRODUCTION

Chapter 3 showed that there does appear to be an impairment in core cognitive functions in mood disorders and Chapter 4 suggested that these impairments may explain wider cognitive dysfunction. Given that associations between brain structure and cognitive impairment have been established, investigating the neural correlates of core cognitive dysfunction is important for improving our understanding of cognitive functioning in mood disorders.

In BD, structural abnormalities are well documented, including abnormal cortical thickness compared to HCs across widespread areas of the brain (Foland-Ross et al., 2011; Kuang et al., 2022; Lan et al., 2014; Necus et al., 2021). Abnormal cortical morphology has been found in euthymia, suggesting that such abnormalities may represent a core feature or consequence of BD (Macoveanu et al., 2021). Brain abnormalities have been detected using other measures, such as cortical and subcortical volume, gyrification, and white matter integrity (Hibar et al., 2016; A. McIntosh et al., 2009; Necus et al., 2019). However, results from a large study by the ENIGMA BD Working Group found widespread cortical thinning in BD compared to HCs, but did not detect differences between BD and HCs in cortical surface area, suggesting that cortical thickness may be a more sensitive measure of structural alterations in BD (Hibar et al., 2018). Several specific regions have been associated with decreased cortical thickness in BD, including the frontal and parietal lobes bilaterally, the anterior cingulate cortex (ACC), and the posterior cingulate cortex (PCC) (Foland-Ross et al., 2011; Hibar et al., 2018; Kuang et al., 2022; Lan et al., 2014). These findings are generally supported by meta-analytic results, which suggests there is some degree of abnormal cortical thickness in BD, however, results are heterogeneous and there is not yet a consensus of which regions, if any, are particularly affected (Hanford et al., 2016; Karantonis et al., 2023; Z. Zhu et al., 2022).

Abnormal cortical thickness appears to be associated with cognitive impairment in BD (Hartberg et al., 2011; Hatton et al., 2013; Kang et al., 2022) and may be more strongly related to cognitive impairment in BD than other morphological metrics such as cortical volume and surface area (Karantonis et al., 2023; Macoveanu et al., 2021; Rodrigue et al., 2018). General cognitive function in BD has been linked to cortical thickness primarily in prefrontal regions,

ACC and PCC (Kang et al., 2022; Knöchel et al., 2016; Macoveanu et al., 2021). However, studies into the association between cortical morphology and cognitive function in BD are scarce (Karantonis, Carruthers, et al., 2021) and findings are not consistent, with some studies failing to find evidence of a relationship (Alonso-Lana et al., 2016; Gutiérrez-Galve et al., 2012).

In terms of core cognitive functioning, few studies focus on the relationship between cortical thickness and impairments in PS, SA, or EF in BD (Karantonis, Carruthers, et al., 2021). Studies on PS tend to focus on associations with white matter integrity and report mixed findings (Krabbendam et al., 2000; Masuda et al., 2020; Poletti et al., 2015). The few studies that measure cortical thickness suggest that performance on DSST, TMT-A, and RT tasks may be related to cortical thickness in BD, even in euthymia (Fears et al., 2015; Oertel-Knöchel et al., 2015; Rodrigue et al., 2018). To the author's knowledge, no studies have aimed to specifically test a link between SA and cortical thickness in BD. Poorer performance on CPTs and other tests of attention has been associated with prefrontal and hippocampal volumes (Sax et al., 1999). Some studies included tests of attention as part of a wider neuropsychological battery and support the finding that poorer attention and SA is related to abnormal cortical volume and thickness (Fears et al., 2015; Hatton et al., 2013). EF has been associated with cortical thickness in frontal and temporal areas in BD, including in euthymia (Abé et al., 2018; Oertel-Knöchel et al., 2015). However, some studies failed to find associations between cortical morphology and core cognitive functions (Gutiérrez-Galve et al., 2012; J. U. Kim et al., 2022; Knöchel et al., 2016), often due to significant results not surviving corrections for multiple comparisons (Fears et al., 2015; J. U. Kim et al., 2022). While the current literature appears to point to a link between core cognitive dysfunction and cortical thickness, there is not yet a consensus on which core cognitive functions and brain areas may be involved.

There is therefore a need for more studies to test whether cortical thickness is related to core cognitive dysfunction in BD. Studies that consider performance on each test individually, rather than creating a composite score for general cognitive performance would be useful in disentangling the specific cognitive functions that may be involved. Given the statistical issues with using traditional univariate approaches (A. R. McIntosh & Mišić, 2013; Rothman, 1990), multivariate methods should be used to test complex brain-cognition associations, whilst considering the association between cognitive domains; a feature often overlooked in the literature. Research is beginning to leverage the power of CCA to investigate brain-cognition associations in mood disorders (Ang et al., 2020; Rodrigue et al., 2018), but studies are scarce,

especially for BD. Existing research has not focussed on core cognitive functioning, and studies often miss tests of PS and attention, or focus on general cognitive functioning instead. We therefore elected to examine the associations between core cognitive performance and cortical thickness in patients with euthymic BD using CCA. The data used here were from a wider study into lithium in BD (the Bipolar Lithium Imaging and Spectroscopy Study [BLISS]; Necus et al., 2019, 2021; Smith et al., 2018). All participants underwent structural magnetic resonance imaging (MRI) and completed a battery of neuropsychological tests. We hypothesised that abnormalities in cortical thickness would be associated with poor cognitive performance in patients with BD.

5.2. METHODS

To assess brain-cognition relationships in BD, ideally a large sample (e.g., *n*=100) of BD patients and the same number of healthy controls would be recruited and matched on age, sex, years of education, and premorbid IQ. Some have argued that CCA requires a sample size of at least 20 times the number of variables in the analysis (Dattalo, 2014), however, with the number of variables in an ideal study design, this would require a sample size of over a thousand patients, which is rarely practically possible. Others stated that a sample size of fifty is sufficient to detect strong canonical correlations (Barcikowski & Stevens, 1975). A comprehensive neuropsychological battery for this study should include a variety of tasks recommended by the ISBD Targeting Cognition Task Force (Miskowiak, Burdick, et al., 2017; Yatham et al., 2010), as specified in section 4.2. For example: DSST, TMT-A, and simple-RT tasks to measure PS; CPT to measure SA; and several tests of EF (e.g., digit span backwards, TMT-B, the Stroop task, and verbal fluency tasks). Wider cognitive functions such as verbal memory (RAVLT) and visuo-spatial memory (spatial recognition, pattern recognition, and spatial span tasks) should also be included. Where possible, statistical process should be used to disentangle cognitive processes from neuropsychological test scores to improve task validity. For example, subtracting the symbol copy variant from the original DSST, and subtracting scores from simple parts of EF tasks (e.g., TMT-A, Stroop simple trials, and digit span forwards) from more complex parts of EF tasks (e.g., TMT-B score, Stroop colour-word trials, and digit span backwards, respectively). Ex-Gaussian models should be applied to CPT data to disentangle elements of PS from SA. Structural MRI scans should be performed using the same scanning protocol for each participant.

Due to external factors, data collection was not possible for this thesis, therefore, we chose extant datasets that matched this design as closely as possible. We used a dataset that was collected as part of the Bipolar Lithium Imaging and Spectroscopy Study (BLISS; Smith et al., 2018). T1-weighted imaging, diffusion-weighted imaging and lithium imaging (^{7}Li) data have been reported previously (Necus et al., 2019, 2021; Smith et al., 2018). BLISS was granted a favourable ethical opinion by a United Kingdom National Research Ethics Committee (14/NE/1135) and all participants provided written informed consent.

5.2.1. Participants

Fifty-nine patients with BD were recruited. Inclusion criteria for patients were: a diagnosis of BD (type I or II) according to DSM-5 (American Psychiatric Association, 2013) criteria and euthymic mood at entry and assessment. Diagnosis was confirmed through clinical interview by an experienced research assistant (CF), supervised by a senior psychiatrist (DC) via discussion of the assessment and suitability for inclusion. Comorbid psychiatric diagnoses were permissible (excluding neurodevelopmental disorders, neurocognitive disorders, and substance abuse) if the primary diagnosis was BD, confirmed by a senior psychiatrist (DC) reviewing case notes as required. Euthymia was defined as a score of <7 on both the YMRS (Young et al., 1978) and the 21-item Hamilton Depression Rating Scale (HAM-D; Hamilton, 1960). Twenty-nine BD patients were taking lithium; the other patients were taking maintenance treatments but were naïve to lithium.

Twenty-eight HCs were recruited. HCs had no history of psychiatric illness and were not taking any psychotropic medications. All subjects were 18-65 years of age. Exclusion criteria for all subjects were: contraindications to MR examination (e.g., cardiac pacemaker claustrophobia, exceeding 150kg in weight), current or past medical condition likely to affect brain structure; current or recent substance abuse (NetSCID Module E); alcohol intake exceeding 21 units per week (self-reported); and a learning disability or impairment of capacity.

5.2.2. Materials

5.2.2.1. Clinical measures

To confirm diagnosis and assess psychiatric comorbidities, the NetSCID diagnostic tool was used; this is a validated online version of the Structured Clinical Interview for DSM-5 criteria (SCID; Telesage, Inc., Chapel Hill, NC, USA). The Young Mania Rating Scale (YMRS; Young et al.,

1978) was used to measure current manic symptoms. The 21-item Hamilton Depression Rating Scale (HAM-D; Hamilton, 1960) was used to measure current depressive symptoms.

5.2.2.2. Neuropsychological tests

A battery of neuropsychological tests was used to measure core cognitive functions.

5.2.2.2.1. d2 test of attention

The d2 test of attention is a paper-and-pencil cancellation test of attention and concentration processes (Bates & Lemay, 2004). During the test, participants were simultaneously presented with many visual stimuli on one page that are similar (the letters 'd' and 'p' with one to four dashes placed anywhere above or below each letter). There are 14 rows of stimuli, with each row having 47 letters. The target stimulus was the letter 'd' with two dashes above or below it (hence 'd2'), and the non-target distractors were either the letter 'd' with more or less than two dashes above/below it or the letter 'p' with any number of dashes above/below it. Participants were instructed to cancel out as many targets as possible from left to right in one row within 20 seconds. This was repeated 14 times (i.e., one row is completed for each trial) without pauses between trials. Three outcome measures were obtained from this task: the percentage of correct responses; concentration performance (total number of correctly cancelled minus total number incorrectly cancelled); and fluctuation rate (maximum total items processed in a trial minus minimum total items processed in a trial).

5.2.2.2.2. The Digit Symbol Substitution Test (DSST)

The DSST subtest of the Wechsler Intelligence Scales (Wechsler, 2011) was used. It is highly sensitive to neuropsychological dysfunction in many disorders and is generally used as a measure of PS (Joy, Fein, et al., 2003; Lezak, 1995). Participants transcribed a unique geometric symbol with a corresponding Arabic number, shown in a key at the top of the page. The number of correct responses in 90 seconds and the time to complete the task were recorded. Participants also completed a sub-test of the DSST, Symbol Copy, which involves copying symbols from one row of boxes to another. This retains the motor speed element but does not require other cognitive processes involved in transcribing. In our study, the outcome measure was the time taken to complete the DSST original version minus the time taken to complete the Symbol Copy version; this removes the motor time from the original score, providing a metric that more accurately reflects cognitive PS without the motor component.

5.2.2.2.3. Trail Making Test (TMT)

The TMT is a pen-and-paper test of attention, PS and EF. In Part A of the task (TMT-A), participants connected a series of numbered circles spread on a page in ascending order (e.g., 1, 2, 3, 4, etc.), and the time taken to complete this task was the raw score. TMT-A therefore captures elements of motor speed, cognitive processing speed, and attention and can be used on its own as a measure of these functions. In Part B of the task (TMT-B), participants connected a series of circles on the page that are either numbered or contain a letter, in alternate ascending and alphabetical order (e.g., 1-A-2-B-3-C, etc.), with time to complete the task as the raw score. TMT-B tests WM and switching (elements of EF), as well as motor speed, cognitive speed, and attention. The latter three functions were fractioned out of the TMT-B score by subtracting the time to complete TMT-A from the time to complete TMT-B; what was left represented the additional time taken to switch between the two series, reflecting the executive control element.

5.2.2.2.4. Category fluency

In category fluency tests, participants were asked to name as many words in one category as possible in a minute. In this study, the animal naming version was used. The outcome measure was the number of correctly names animals in 60 seconds. Verbal fluency likely requires a range of cognitive functions, including PS, attention, WM, and EF, and is often used in the literature as a measure of PS or EF.

5.2.3. Procedure

All subjects attended a screening visit to confirm eligibility and complete structured clinical interview. MRI scans were performed at 9am and participants taking lithium were instructed to take lithium as usual the night before.

5.2.3.1. Scanner and image acquisition

MR data were acquired on a Philips 3-T Achieva scanner (Philips Medical Systems) with 1H structural and diffusion weighted imaging performed using a Philips 8-channel SENSE head coil. T1-weighted images (T1w) and T2-weighted (T2w) Fluid-Attenuated Inversion Recovery (FLAIR) images of brain anatomy were acquired for each subject with 1H gradient echo sequence (TR=9.6 ms; TE=4.6 ms, FOV=180 mm^3, acquisition matrix 240 x 208 x 180 mm^3, 1 x 1.15 x 1 mm^3, reconstructed into a matrix size of 180 x 256 x 256, 1mm^3 average).

5.2.4. Data analysis

5.2.4.1. MRI data pre-processing

All images were exported in DICOM format and converted to NIFTI format using the dcm2niix program (Li et al., 2016). Data pre-processing and analysis were performed using FreeSurfer 6.0 (Fischl, 2012), utilising the standard *recon-all* pipeline, including removal of non-brain tissue, segmentation of grey matter and white matter surfaces, and cortical parcellation. FLAIR images were used in the *recon-all* pipeline to improve the pial surface and segmentation. Following the *recon-all* pipeline, cortical thickness measures were obtained for each anatomical ROI for each subject. *n*=68 ROIs were defined using the Desikan-Killiany atlas (Desikan et al., 2006).

Subjects were excluded if the T1-weighted images included MR artefacts or were poor quality (*n*=1; HC). For the remaining subjects, visual quality inspections were performed on a sub-sample of participants (*n*=47 [54%] of total sample; *n*=30 BD and *n*=18 HC) in FreeSurfer's FreeView function. This sub-sample included all participants whose data were considered an outlier (defined as ±3 SD from the mean) in one or more ROIs, as well as *n*=4 randomly sampled participants with no outliers in any ROIs. Manual corrections of the pial surface were conducted where necessary using FreeView. The pial surface of *n*=24 subjects (i.e., 51% of the subjects who were checked) were edited. After manual corrections, the overall image quality and FreeSurfer processing was deemed to be adequate. Cortical thickness data for each ROI were then extracted for each participant.

ROIs pertaining to the cerebellum were removed as adequate images of this region were not acquired for all subjects. Cortical thickness data for the remaining ROIs were log transformed, then age and sex were regressed out in the control group using robust regression for each ROI. Any ROIs that had skewed residuals following this regression were transformed using Box Cox transformations (Box & Cox, 1964) and then regressions were run on the transformed data. Box Cox transformations of six ROIs were required for the control group (left entorhinal, inferior temporal, precentral, superior parietal, temporal pole areas and right frontal pole), and the same transformations were applied to the BD group. Age and sex were then regressed out of the BD group data using the regression model that was estimated using the control group. Residuals following age and sex regression were then standardised (z-scored) based on the control group mean and SD for each ROI.

5.2.4.2. Neuropsychological data pre-processing

Some subjects had missing data for NART IQ and neuropsychological variables: 5% of data were missing in total (BD=4%, HC=8%). Missing data were spread across neuropsychological variables as follows: NART IQ 7%, d2 test of attention 2%, DSST 7%, TMT-A 7%, TMT-B 8%, and verbal fluency 6%. The pattern of missing data was not considered biased (i.e., missing cases were not influenced by cognitive ability and were 'Missing at Random') and therefore was suitable for imputation. Missing data were imputed using linear regression, where missing data for each variable was predicted using age, education, NART, and all other neuropsychological scores. This was done for BD and HC groups separately since group differences were expected. Before imputing missing data, outliers (±3 standard deviations around the mean) were Winsorized and skewed data were transformed to meet the assumptions of multiple regression, then the imputed values were back-transformed to align with the original data. Variables were reversed where appropriate so that a higher score reflects better performance.

Age and pre-morbid IQ (NART) were regressed out of the complete neuropsychological data (i.e., after data imputation) in the control group using robust regression. All residuals were normally distributed for the HCs. Age and pre-morbid IQ were then regressed out of the BD group data using the regression model estimated for the control group. Residuals following age and pre-morbid IQ regression were standardised (z-scored) based on the HC group mean and SD. The neuropsychological variables were then transformed as required to fit assumptions of multivariate normality and remove outliers: d2 correct, d2 fluctuation rate and DSST were transformed using reflect and log; TMT-A and TMT-B were transformed using reflect and square root; and d2 continuous performance and category fluency did not require transformation. No transformations were required for the HC data. The subsequent analysis was run using the untransformed data as well to compare results.

5.2.4.3. Statistical analysis

Statistical analysis was performed using R Studio version 4.1.2 (R Core Team, 2021). Unless stated otherwise, tests were regarded as significant if $p<0.05$. As well as the HC subject excluded from the analysis due to an MRI artefact, four subjects (n=2 BD, n=2 HC) were removed from the analysis due to a substantial amount of missing cognitive data.

5.2.4.3.1. Group comparisons and ROI analysis

Prior to statistical analysis, continuous variables were tested for normality of distribution using the Shapiro-Wilk tests and by examining Q-Q plots. Student's t-test was used to compare groups where data were normally distributed, otherwise the non-parametric Mann-Whitney U test was used. Dichotomous data were analysed using a $\chi 2$ test.

5.2.4.3.2. Canonical Correlation Analysis (CCA)

CCA was used to assess multivariate associations between neuropsychological performance and cortical thickness of regions, accounting for the covariance within each dataset. CCA is useful when there are high intercorrelations within variable sets (Lambert et al., 1988), which makes it appropriate here, since we expected the cognitive and the cortical thickness variables to be interrelated within their own datasets. Since the cortical thickness and neuropsychological data were both z-scored based on control means and SDs, the data used in the CCA reflected abnormalities compared to HCs. We therefore performed CCA in the BD group only, to assess which abnormalities in cognitive functioning were associated with abnormalities in cortical thickness.

The number of subjects in each group was less than the number of ROIs, so PCA was applied to the cortical thickness data first to reduce dimensionality (A. R. McIntosh & Mišić, 2013; H. T. Wang et al., 2020). This step has been recommended to limit the feature to sample ratio and avoid overfitting (Dinga et al., 2019; Mihalik et al., 2022). The number of principal components (PCs) that collectively explained 90% of this dataset were retained and these PCs were then fed into the CCA as the cortical thickness dataset. The CCA therefore tested the association between the seven neuropsychological variables and the PCs of the cortical thickness dataset. The contribution of each original variable to the first canonical variate was estimated by correlating it with the canonical variate (i.e. canonical loadings), where loadings above 0.3 or below -0.3 were considered non-trivial (A. R. McIntosh & Mišić, 2013).

As well as interpreting the strength of the canonical correlations, their statistical significance was tested in two ways. Firstly, Pillai's trace was used as a test of significance; this was chosen over other options (e.g., Wilk's test) as it is considered the most conservative and robust test of all the multivariate *F*-tests of significance (Dattalo, 2014; Pillai, 1955). Secondly, a permutation test was used to verify the results of Pillai's trace (Bullmore et al., 1999; A. R. McIntosh & Mišić, 2013). In the permutation test, the order of subjects (rows) in the cortical

thickness dataset (*V*) was randomly shuffled so that it no longer aligned with the cognitive dataset (*U*), and then CCA was performed on these mismatched datasets. This was done 10,000 times to produce a distribution of statistics that were produced using the randomly shuffled data. We would expect a true significant result of the original CCA analysis to have a more extreme value than the statistics from the randomly shuffled data. To account for some extreme statistics being produced in the randomly shuffled data by chance, a rule of thumb is to allow for 5% of the permuted results to be more extreme than the original results. In this sense, the permutation test produces a *p*-value that represents the percentage of randomly permuted results that were more extreme than the original CCA. This *p*-value can be interpreted in the same way as usual, with a significant result being inferred from a $p<.05$.

5.3. RESULTS

5.3.1. Group differences

Table 15 shows the demographic information and clinical characteristics for patients with BD and HCs. Groups did not differ in age or sex. HCs had significantly more years of education than BD. BD scored significantly higher on HAM-17 and YMRS than HCs. Groups did not differ in premorbid IQ (NART score) or on average cortical thickness in the left or right hemispheres. Table 15 also displays neuropsychological performance and group differences. The BD group showed significantly poorer neuropsychological performance than HCs on most tests. The group differences were also tested after accounting for age and premorbid IQ and imputing missing values; Figure 40 illustrates these z-scores for each neuropsychological test and the group differences. The pattern of results was the same as above, except TMT-B was not significant ($p=.063$, $d=0.51$). Table 33 in Appendix C shows the full results of the group differences after regressing out age and premorbid IQ and imputing missing values.

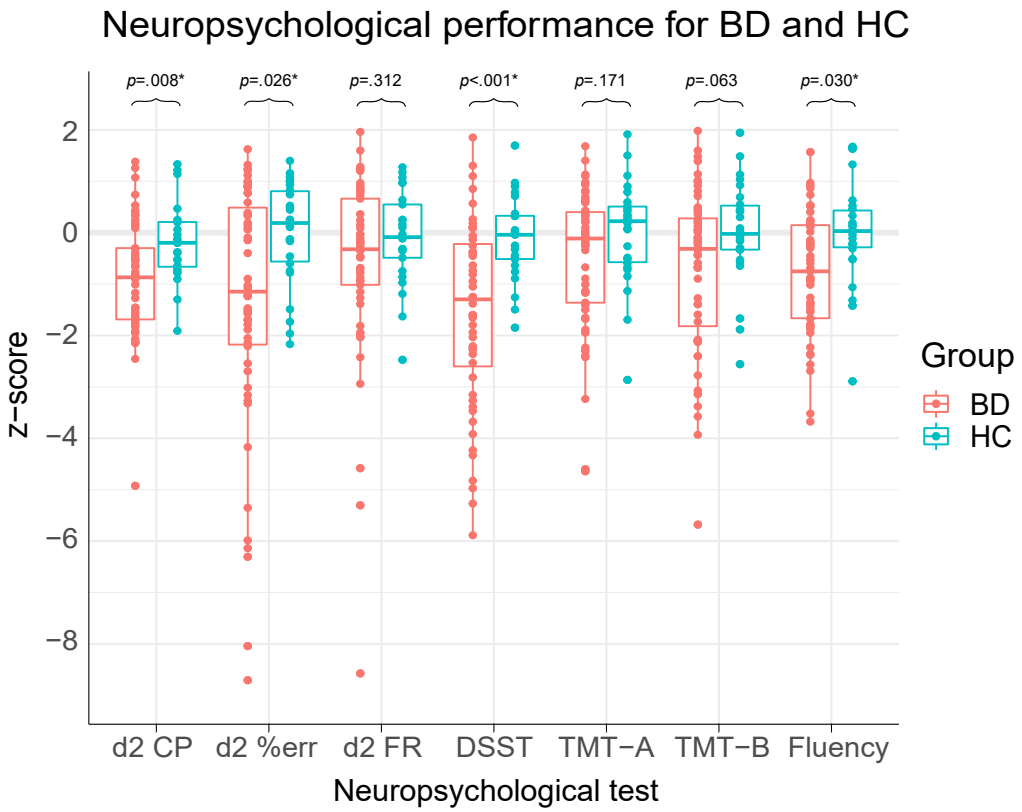

Figure 40: Neuropsychological performance for bipolar disorder (BD) and healthy control (HC) groups. Z-scores were based on control group mean and standard deviation. CP=concentration performance; %err=percentage of errors; FR=fluctuation rate, DSST=Digit Symbol Substitution Test; TMT=Trail-Making Test.

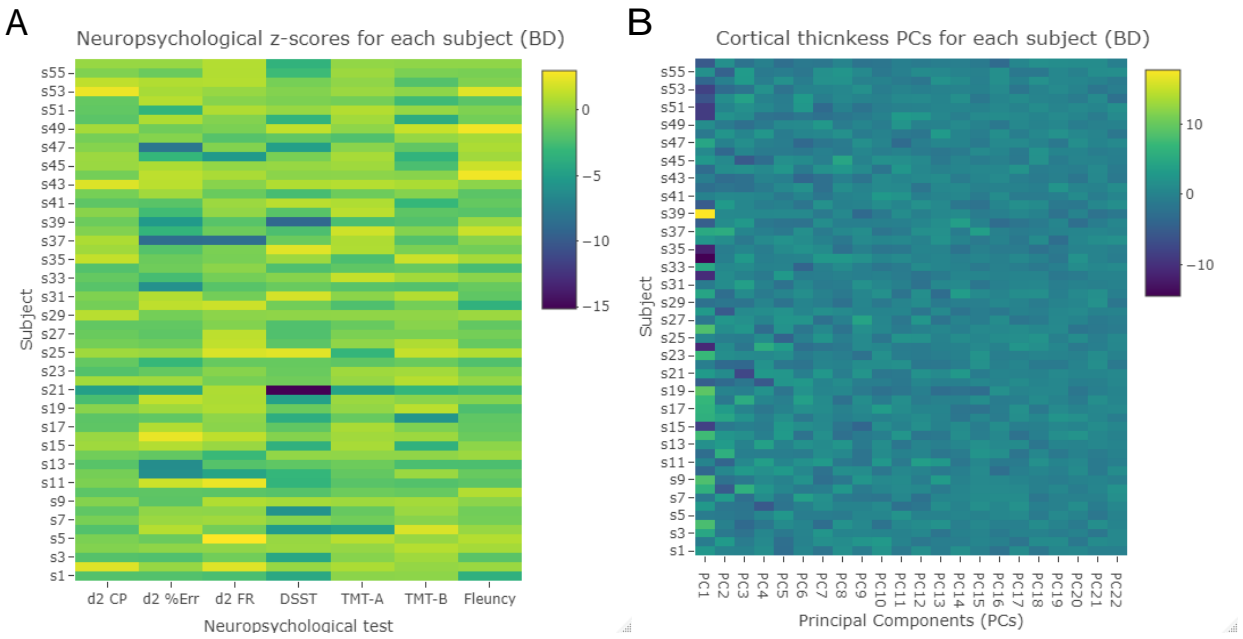

Figure 41: Heatmaps of neuropsychological and cortical thickness data for each BD subject. A) Neuropsychological data z-scored based on control group mean and standard deviation (dataset U). B) Cortical thickness PCs (n=22 explained 90% of the cortical thickness data; dataset V). BD=bipolar disorder; PC=principal component; CP=concentration performance; %Err=percentage of errors; FR=fluctuation rate; DSST=Digit Symbol Substitution Test; TMT-A=Trail-Making Test.

	Healthy Control group				Bipolar Disorder group				Group differences			
	N	**Mean**	**SD**	**Median**	**N**	**Mean**	**SD**	**Median**	**Statistic (*t*/*W*)**	**df**	**p-value**	**Cohen's *d***
Demographics and clinical characteristics												
Sex (N female, % female)	26	12	0.46	-	56	*n*=35	62%	-	1.33	1	.249	-
Age	26	48.46	11.80	48.50	56	45.36	12.15	46.00	612.5	-	.252	0.261
Education (years)	26	16.46	3.78	16.00	56	14.14	2.88	13.00	470.0	-	.009*	0.737
HAM-17	26	1.00	1.57	0	56	4.20	4.16	3.00	1123.5	-	<.001*	0.909
YMRS	26	0.12	0.43	0	56	1.06	2.23	0	938.0	-	.007*	0.513
Medication (N lithium, % lithium)	-	-	-	-	56	*n*=29	52%	-	-	-	-	-
Neuropsychological data												
NART IQ	24	15.54	7.54	14.50	52	17.48	6.76	18.00	1.08	40.70	.288	0.280
d2 Concentration Performance	25	177.92	33.41	177.00	55	154.15	45.12	144.00	439.5	-	.010*	0.575
d2 Percent Error	25	0.03	0.02	0.03	55	0.05	0.05	0.04	876.5	-	.050	0.566
d2 Fluctuation Rate	25	10.64	2.78	11.00	55	12.11	4.99	12.00	809.5	-	.205	0.336
DSST Symbol minus Copy time	23	73.12	13.28	74.50	53	93.03	35.33	88.55	856.0	-	.005*	0.662
TMT-A time	23	27.71	11.90	26.55	54	30.64	11.31	28.37	725.0	-	.249	0.258
TMT-B minus TMT-A time	23	27.52	16.96	23.40	52	39.16	24.02	32.09	770.0	-	.049*	0.533
Category Fluency total	23	24.96	4.96	25.00	54	21.80	6.24	21.50	-2.36	51.85	.022*	0.544
Cortical thickness												
Average left hemisphere	26	2.46	0.10	2.45	56	2.43	0.09	2.42	-1.70	46.82	.095	0.416
Average right hemisphere	26	2.46	0.09	2.45	56	2.43	0.10	2.41	-1.28	50.23	.206	0.304

*Table 15: Descriptive statistics and group differences on demographics, clinical characteristics, neuropsychological variables, and structural MRI variables. *Significant at the 0.05 level. SD=standard deviation; df=degrees of freedom; HAM-17= Hamilton Rating Scale for Depression 17-item; YMRS=Young Mania Rating Scale; NART=National Adult Reading Test; DSST=Digit Symbol Substitution Test; TMT=Trail-Making Test.*

5.3.2. Canonical Correlation Analysis (CCA)

CCA assumes that the variables have linear relationships with one another and the two datasets have multivariate normality (Sherry & Henson, 2005). Scatterplots were visually inspected to confirm linear relationships between the variables. The neuropsychological data did not have multivariate normality for the BD group, according to Mardia's tests: skewness $p<.001$ and kurtosis $p=.007$. Therefore, neuropsychological data were transformed using reflect and log or reflect and square root transformations where appropriate. After transforming, data showed multivariate normality ($p>.05$ for skewness and kurtosis). The subsequent analysis was run using the transformed data; we also ran the analysis using the untransformed data for comparison. PCA was conducted on the cortical thickness of ROIs of the BD group to reduce data dimensionality. 22 PCs explained 90% of the data and were retained for the CCA. The 22 PCs had multivariate normality ($p>.05$ for skewness and kurtosis). Figure 41 shows heatmaps of the standardised neuropsychological data and the cortical thickness PCs for BD participants.

CCA was performed using the 7 neuropsychological variables as dataset *U* and the 22 cortical thickness PCs at dataset *V*. Figure 42 shows the first canonical correlation coefficient, which had a strong linear correlation of .832. Pillai's Trace statistic showed that the first canonical variate was significant ($p=.043$). The remaining canonical correlations were as follows: 0.779, 0.727, 0.706, 0.655, 0.455, and 0.340; the first three were significant ($p<.05$). A permutation test with $n=10{,}000$ permutations was run to test the robustness of the first canonical correlation. The permutation test suggested the correlation was not significant but was trending towards significance ($p=.085$).

Table 16 displays the canonical loadings cross-loadings for each cognitive variable from set *U*. d2 fluctuation rate, DSST, and TMT-B were more strongly associated with the first cognitive canonical variate *V1* (cross loadings were -0.322, -0.306, and 0.374, respectively). Table 17 displays the ROIs that were associated with the first canonical variate. Figure 43 illustrates the association between cortical thickness in each region and the cognitive dataset. The canonical cross-loadings suggest that the regions associated with the first canonical variate *U1* were: the left posterior cingulate (PCC), left superior temporal, right parahippocampal, right entorhinal and right lateral occipital areas (all cross loadings > 0.3). The left parahippocampal, right PCC

and right superior temporal regions appeared to have a significant association (p<.05) but did not meet the threshold of 0.3.

When the analysis was applied to the untransformed neuropsychological data, the first canonical correlation showed a strong linear correlation of .824 but this was not significant according to Pillai's Trace (p=.07) or the permutation test (p=.13). Canonical cross-loadings suggested that that d2 percent error, d2 fluctuation rate, and DSST scores were associated with cortical thickness in the following areas: paracentral area bilaterally; PCC and superior parietal areas in the left hemisphere; and lateral occipital and lingual areas in the right hemisphere (all coefficients ±0.3).

Variables from *U*	**Loadings**	**Cross-loadings**
d2 Continuous Performance	-.059	-.049
d2 Percentage of Errors	-.227	-.189
d2 Fluctuation Rate	-.387	-.322*
DSST	-.368	-.306*
TMT-A	.299	.249
TMT-B	.450	.374*
Category Fluency	-.163	-.136

*Table 16: Canonical loadings and cross-loadings for each variable in the cognitive dataset (U). *Associated with the first canonical variate V1 at the 0.3 threshold.*

	Left hemisphere		**Right hemisphere**	
Region	***rho***	***p***	***rho***	***p***
Entorhinal	.224	.098	.360*	.006
Lateral occipital	.150	.271	.316*	.018
Parahippocampal	.283	.035	.302*	.024
Posterior cingulate	.312*	.019	.287	.032
Superior temporal	.359*	.007	.279	.038

*Table 17: Brain regions associated with the first canonical variate U1. *Associated with the first canonical variate U1 at the 0.3 threshold.*

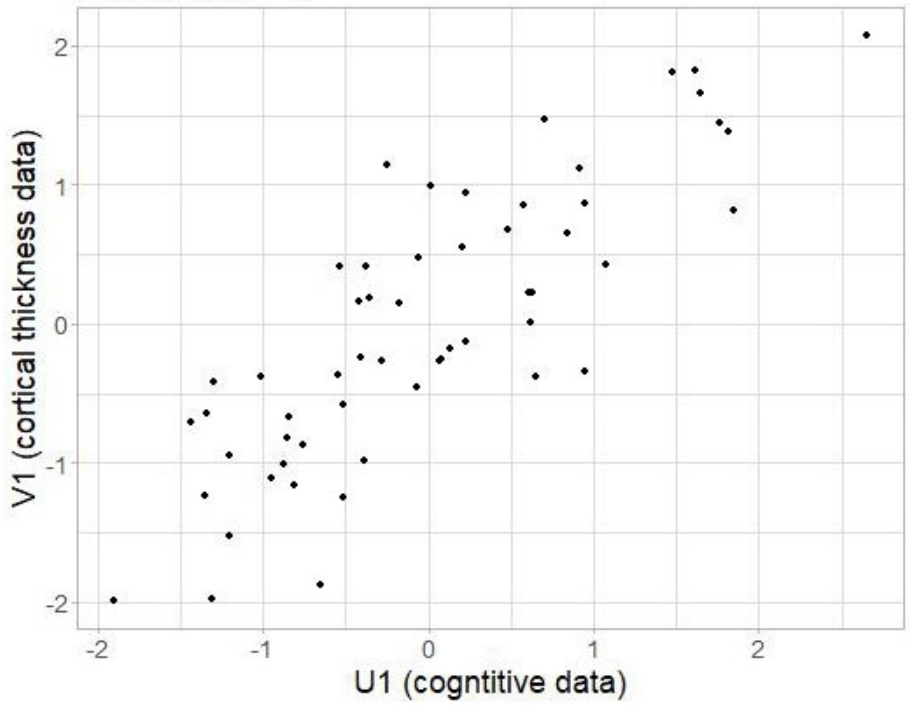

Figure 42: Scatterplot of the correlations between the first canonical variate from each dataset, U1 and V1.

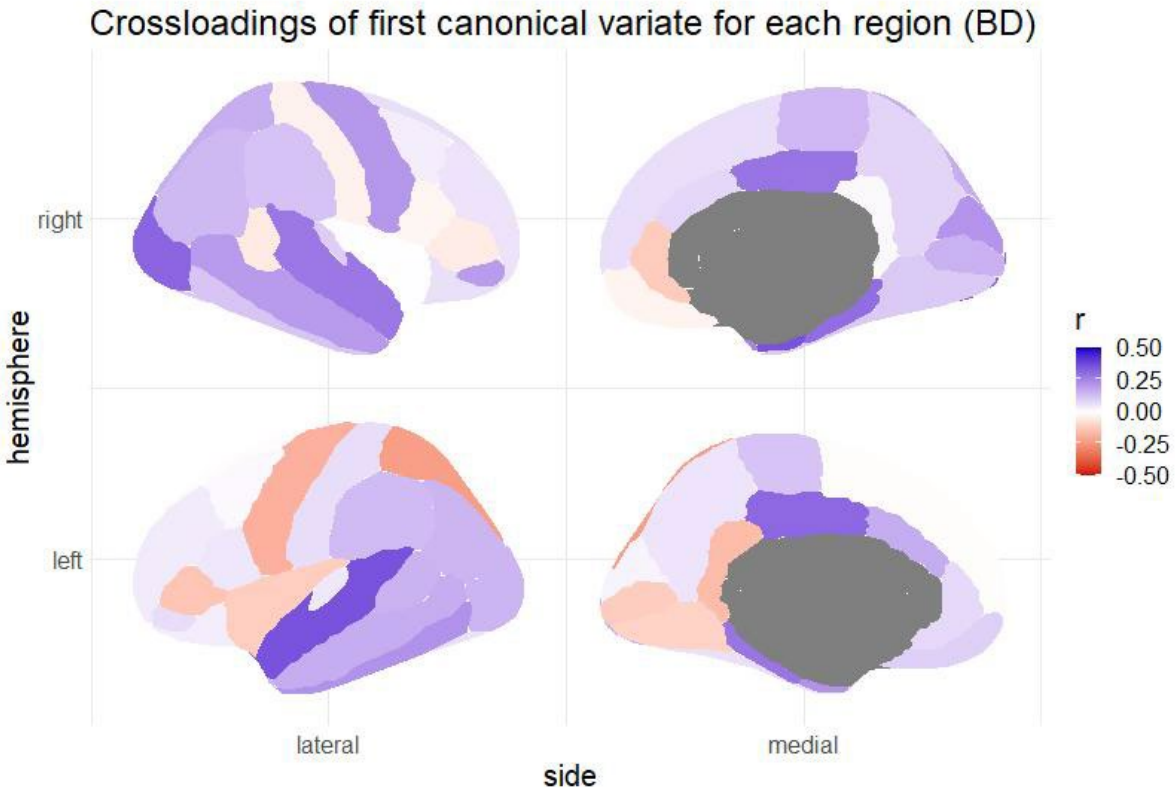

Figure 43: Canonical cross-loadings illustrating the association between cortical thickness in each region and the first canonical variate from the cognitive data (U1).

5.4. DISCUSSION

5.4.1. Summary of results

The euthymic BD group showed poorer performance than HCs on some tests of PS, EF, and attention, even after controlling for age and premorbid IQ. Core cognitive impairments were associated with structural brain abnormalities in our BD sample according to Pillai's trace. However, the CCA result did not survive a permutation test. Specifically, poorer scores on d2 fluctuation rate, DSST, and TMT-B were associated with abnormal cortical thickness. The PCC, superior temporal, parahippocampal, right entorhinal and right lateral occipital areas were particularly associated with this impairment.

5.4.2. Comparison with previous literature

Our results are consistent with previous research showing core cognitive impairments in BD (Bo et al., 2017; Cardenas et al., 2016). We did not find an impairment in TMT-A in our BD sample, whereas previous research suggested that TMT-A is one of the most robust measures of cognitive impairment in euthymic BD (Cardenas et al., 2016). This discrepancy may be explained by our sample size and sources of heterogeneity that we did not control for here, e.g., illness severity and psychosis (Bora et al., 2007; Bourne et al., 2013).

Our results are also consistent with previous research that found associations between general cognitive ability and brain structure in BD, including thicker cortex, larger volumes, and smaller area in mostly frontal and parietal regions (Rodrigue et al., 2018). Other research has suggested that structural abnormalities in fronto-limbic areas are related to poor cognitive and general functioning in BD (Dusi et al., 2019). Our results are in line with this: poorer cognitive performance was related to more abnormal cortical thickness in several regions including limbic areas (PCC, entorhinal, and parahippocampal areas). The association of abnormal PCC structure with cognitive impairment in BD is consistent with previous research (Kang et al., 2022; Knöchel et al., 2016). Kang et al. (2022) found that general cognitive performance was related to cortical thickness in BD in several regions, including the PCC and superior temporal areas. Fears et al. (2015) more specifically linked PS (DSST) with cortical thickness of the PCC, however, this result did not survive corrections for multiple comparisons and no other cognitive tests was related to the PCC. We found that the PCC was related to performance on the DSST, but also a test of attention and EF. While Fears et al. used linear regression models, we used multivariate methods, which avoided the problem of multiple

comparisons and allowed us to consider cognitive functions together; this perhaps explains the differences in our results.

Associations between cognitive impairment and abnormal PCC, parahippocampal and entorhinal regions may relate to abnormal default mode network (DMN) functioning in BD. The DMN is a network of brain regions that are active during resting states (Mazoyer et al., 2001; Raichle et al., 2001). The PCC is an important hub within this network (Alves et al., 2019; Leech & Sharp, 2014; Mazoyer et al., 2001); parahippocampal and entorhinal regions are also involved in the DMN (Buckner et al., 2008). The DMN is thought to be related to cognitive functioning, whereby less deactivation of the DMN during tasks causes attentional lapses (D. H. Weissman et al., 2006). The PCC in particular is thought to modulate interactions of the DMN and cognitive-control networks to control allocation of attention (Leech et al., 2011). Hypoconnectivity of the DMN has previously been linked to mood disorders (Meda et al., 2014; Renner et al., 2017) and BD show less variable temporal connectivity strength between regions in the DMN, particularly the PCC, compared to HCs, suggesting a state of rigid connectivity (Nguyen et al., 2017; Rashid et al., 2014). Abnormal activity of the DMN has been linked to cognitive performance in BD, with patients showing a failure to deactivate the DMN during tasks (Miskowiak & Petersen, 2019; Zarp Petersen et al., 2022). Less variable connectivity within the DMN has also been linked to slower PS and poorer EF in BD (Nguyen et al., 2017). The DMN may therefore have a role in core cognitive impairment in BD, which should be further explored.

Concerning the direction of the relationship between cortical thickness and neuropsychological impairment in BD, our results were mixed: d2 test of attention, DSST, and category fluency had negative relationships with cortical thickness of the regions that were more strongly related to cognitive functioning, whereas TMT scores had positive relationships. Thinner cortex is generally assumed to be related to poorer cognitive functioning (Hanford et al., 2016; Kang et al., 2022; Knöchel et al., 2016), however, some studies found that increased, rather than decreased, cortical thickness was related to cognitive impairment in BD (Karantonis, Rossell, et al., 2021; Macoveanu et al., 2021). It could be that for some regions, a thinner cortex is associated with cognitive impairment, while for others the opposite may be true. It may then be prudent to consider only 'cortical abnormalities' in patients compared to controls, rather than discussing the direction of the relationship. Here, we analysed cortical thickness in BD relative to HCs, conceptualising the data as abnormality from the norm, and

discussed results in terms of 'cortical abnormalities'. More research is needed to decipher whether cortical thinning or thicker cortex may be related to cognitive impairment in BD, and whether this effect is similar across the brain or whether it changes depending on the region.

5.4.3. Limitations

Several limitations of this study should be considered. Firstly, the neuropsychological test battery was limited; it included measures of core cognitive functions but did not include other functions such as WM, or verbal and visual learning and memory. The measure of attention did not specifically measure SA. SA has been arguably understudied in mood disorders and may be of particular interest here due to its relationship with the DMN (D. H. Weissman et al., 2006). Future research should employ tests of SA to further investigate this link.

Secondly, our sample size was modest: some have argued that CCA requires a sample size at least 20 times the number of variables in the analysis to adequately estimate the first canonical pair (Dattalo, 2014). Others stated that a sample size of 50 is sufficient to detect strong canonical correlations (>.7; Barcikowski & Stevens, 1975). Implementing PCA before CCA mitigates this issue to some extent (Mihalik et al., 2022), nevertheless, our results may have been different with a larger sample. Further, we did not perform an additional cross-validation step, which is often advised to validate the results of CCA (Dattalo, 2014; Dinga et al., 2019). Our sample was arguably too small to split into training and tests datasets to perform cross-validation; our study should therefore be replicated with a larger sample.

CCA is correlational, so we cannot infer any cause-effect relationships. Other limitations of CCA include overfitting, which has been demonstrated in a study of patients with depression (Dinga et al., 2019; Drysdale et al., 2017), therefore our results should be replicated. Our first canonical correlation did not survive a permutation test, therefore our interpretation of the brain-cognition associations detected here may not reflect true associations, or our analysis was perhaps underpowered. Dinga et al. (2019) similarly found that a strong first canonical correlation did not survive a permutation test. Their 10-fold cross validation of CCA results suggested that results may not be reproduceable in different datasets. In our study, the first canonical correlation did not survive permutation test, so the subsequent canonical correlations were not interpreted. In the CCA by Rodrigue et al. (2018), the second canonical pair indicated brain-cognition associations distinct to results from the first canonical pair. Assessment of subsequent canonical pairs may be useful in future research utilising CCA.

An ongoing problem in researching the cognitive profile of BD is heterogeneity within the disorder: BD groups tend to vary in the degree of cognitive impairment, with some patients performing as well as HCs (Cullen et al., 2016). Some evidence points towards the existence of subgroups with distinct cognitive profiles in BD (Chakrabarty et al., 2021; J. Jensen et al., 2016). We did not take this into account in our study, however we did find that BD varied in cognitive impairment. Patients with BD also vary in brain structural abnormalities (Nunes et al., 2020), with some studies similarly suggesting distinct subgroups of patients based on cortical thickness, cortical volume, and surface area (Doan et al., 2017). A recent review of 20 studies did not find strong evidence to support the idea that cognitive subgroups of BD map onto distinct structural brain abnormality profiles (Karantonis et al., 2023; Karantonis, Rossell, et al., 2021); our approach may therefore be justified in light of these results.

Clinical confounds such as illness duration, illness severity, and psychosis, which have been shown to be related to cognitive impairment and brain abnormalities in BD were not controlled for here (Cullen et al., 2016; Foland-Ross et al., 2011; Hanford et al., 2016; Hibar et al., 2018). Some studies have found brain-cognition associations even after controlling for clinical confounds (Macoveanu et al., 2021). It may therefore be possible that the associations we found are independent of clinical characteristics, but further research should confirm this. Similarly, childhood trauma has been associated with both cognitive impairment and brain structural abnormalities in mood disorders (Bücker et al., 2013; Jørgensen et al., 2023), therefore this may be a confounding variable in brain-cognition analyses. We controlled for age to account for the decline of both cognitive functioning and brain structure with age (Zimmerman et al., 2006). Our sample was euthymic at the time of testing, so we could not assess the effect of mood state on brain-cognition associations. Other studies found that euthymic and depressed BD patients may differ in cortical thickness in some areas (Kang et al., 2022), so future research should investigate the effect of mood.

Here, we used cortical thickness to measure structural abnormalities in BD. Previous research suggested that this may be a more sensitive measure of structural alterations in BD, and more strongly related to cognitive impairment in BD, than other structural metrics such as cortical volume and surface area (Hibar et al., 2018; Karantonis et al., 2023; Macoveanu et al., 2021; Rodrigue et al., 2018). However, cortical volume, surface area, and gyrification have nevertheless been associated with core cognitive functioning in BD, so should not be discounted in future studies (Abé et al., 2018; Fears et al., 2015; A. McIntosh et al., 2009;

Oertel-Knöchel et al., 2015; Sax et al., 1999). Future research should similarly consider investigating associations between core cognitive impairment and structural and functional connectivity, which have previously been related to cognitive functioning in BD (Deng et al., 2019; Macoveanu et al., 2021; Yu et al., 2021). We also acknowledge that we did not analyse subcortical regions that are known to be involved in cognition in BD, such as the hippocampus, amygdala, and cerebellum (Hartberg *et al.*, 2011; Miskowiak *et al.*, 2016; McPhilemy *et al.*, 2020).

5.4.4. Future directions

Another limitation with measuring cortical thickness concerns the mixed findings of whether poorer cognitive functioning is related to thinner or thicker cortex. The direction of the relationship between cortical thickness and cognitive functioning depends on other dimensions of cortical morphology: cortical thickness, volume, and surface area have been shown to covary in the healthy human brain according to a universal scaling law (Y. Wang et al., 2016, 2020). Wang et al. (2020) developed a more sophisticated model of brain morphology than the traditional metrics, that accounts for covariance between individual features and provides metrics that can be interpreted in a similar way to the traditional measures. A demonstration of this model detected brain morphological abnormalities in Temporal Lobe Epilepsy over and above the typical changes seen in healthy ageing (Y. Wang et al., 2020). Future research should utilise these independent components of cortical morphology to more precisely assess neural correlates of cognitive impairment in BD, accounting for the covariance between cortical thickness, volume, and surface area.

5.4.5. Interim summary

This study was one of few to investigate multivariate brain-cognition associations in BD. The results suggest that impairments in PS, attention, and EF may be associated with cortical abnormalities in euthymic BD. The association may be specific to particular tests of core cognitive functions. The brain regions associated with abnormal cognitive functioning in BD included regions associated with the DMN, suggesting that this may be an avenue for future research to explore. Future research should test these associations using independent components of cortical morphology and should investigate the role of subcortical areas, as well as structural and functional connectivity.

Chapter 6 DISCUSSION

This thesis set out to explore the nature of core cognitive impairments and their association with wider cognitive functioning and brain structure in mood disorders. The following chapter will provide a summary of the empirical results (section 6.1), as well as a discussion of their implications (section 6.2), strengths and limitations (section 6.3), recommendations for future research (section 6.4), and a conclusion of the current body of work (section 6.5).

6.1. SUMMARY OF FINDINGS

In Chapter 3, we systematically reviewed the literature to establish the presence and magnitude of impairments in PS and SA in people with BD and MDD. We meta-analysed data for each neuropsychological score separately to avoid heterogeneity in the results and to investigate the sensitivity of each test to cognitive dysfunction in mood disorders. Where possible, we performed subgroup analysis based on mood state to assess impairments in symptomatic and euthymic/remitted states. BD and MDD showed impairments in PS and SA in most neuropsychological tests. Impairments were found in symptomatic states and in euthymia/remission, however some outcome measures were not sensitive to cognitive impairment in euthymia/remission. Some outcome variables did not show significant impairments overall, but these appeared to be driven by groups of studies with euthymic/remitted groups. Our review highlighted a lack of data for depressed and manic BD groups.

Chapter 4 examined the role of core cognitive functions in wider cognitive functioning in people with BD and MDD. Specifically, we used hierarchical regression to assess the individual and joint effects of PS, SA, and EF on memory in euthymic BD, BD depression, and MDD. PS and EF appeared to explain memory impairments in euthymic BD, whereas SA appeared to contribute to memory impairments in MDD. Memory did not seem to be impaired in our depressed BD sample; however, SA was associated with memory performance. Our results suggest that PS and EF may be central impairments in euthymic states, whereas SA may play a role in cognitive functioning in depressed states. Due to limitations to our results, further research is needed before conclusions can be drawn about the precise nature of the cognitive profile in each diagnostic group and mood state, however, our results highlight potentially

important relationships between cognitive functions, which justifies further research into a potential cognitive hierarchy of dysfunction in mood disorders.

Chapter 5 investigated associations between abnormal brain structure and core cognitive dysfunction in BD. This dataset contained neuropsychological scores on tests of attention, PS and EF, as well as T1-weighted MRI scans from people with euthymic BD and matched HCs. We conceptualised neuropsychological and cortical thickness data for patients as deviations from the control group mean. Multivariate analysis (CCA) was used to allow us to consider complex associations within and between the neuropsychological and cortical thickness data. Impairments in PS, attention, and EF were associated with abnormal cortical thickness in several regions, including the PCC, superior temporal, parahippocampal, right entorhinal and right lateral occipital areas. Notably, several of these areas are associated with the DMN, which has previously been implicated in BD (Nguyen et al., 2017; Rashid et al., 2014).

6.2. Interpretation and implications of findings

6.2.1. Neuropsychological methodology

The results have implications for research and clinical practice. Firstly, the meta-analysis suggested that not all tests of PS and SA are sensitive to cognitive dysfunction in BD and MDD. Further, the results may depend on the precise outcome measures chosen from each test. Previous research has not reached a consensus on which tests may be most sensitive to cognitive impairment in mood disorders: for example, Cardenas et al. (2016) reported that TMT-A was one of the most robust measures of cognitive impairment in euthymic BD, and identified digit symbol coding and CPTs among a group of instruments that showed the highest magnitude of impairment among patients with BD. Previous studies suggested that DSST but not TMT-A performance mediated the relationship between depression status and verbal and visuospatial memory in MDD (Zaremba et al., 2019). Our meta-analysis suggested that DSST, Stroop tasks, and PS composite scores might be the most sensitive to cognitive impairments in BD groups (i.e., had the largest ESs). PS composite scores and CPT RT variability may be the most sensitive to cognitive impairments in MDD groups. In our brain-cognition analysis, some tests of PS, EF, and attention, but not others, were associated with brain structure. Future research should therefore carefully consider the instrument and scoring method of neuropsychological assessment when measuring cognitive functioning in mood disorder groups.

6.2.2. Cognitive hierarchy

Secondly, our results provide support for the existence of a hierarchy of cognitive dysfunction in people with mood disorders and highlight the need for further research to account for interrelationships between cognitive functions. In our hierarchical regression analysis, the order of entry appeared to matter, which suggests that researchers should carefully consider which cognitive function(s) may have the most influence on wider functioning. In investigations of brain-cognition associations, studies should use multivariate methods to address this issue and explore complex relationships within and between cognitive and brain imaging data. The existence of a cognitive hierarchy also has implications for clinical practice in its potential to inform cognitive remediation therapies (CRTs). CRTs are psychological interventions that aim to improve cognitive functioning and quality of life in patients and have been shown to be effective for affective disorders including MDD and BD (Anaya et al., 2012; Miskowiak, Seeberg, et al., 2022; Strawbridge et al., 2021). CRTs require subjects to practice cognitive tasks, so the finding that some cognitive functions may be particularly impaired and may affect wider cognitive functioning suggests that CRTs may benefit from focussing on the cognitive functions that are primarily impaired. Our results suggest that the cognitive functions that may be useful to target may vary depending on diagnostic group and mood state, such as PS and EF in euthymia and SA in depressed states, but more research is needed to confirm which functions are primarily impaired in each case.

6.2.3. State vs trait

We found that while some impairments in PS and SA persisted in euthymia/remission, others were only present in symptomatic states. Further, the structure of a hierarchy of cognitive dysfunction may vary depending on mood state. Some cognitive impairments may therefore represent trait features of MDD and BD that are present throughout the disorder, whereas others may reflect state-dependent functioning. While state-related cognitive impairments may be a consequence of current depressed or manic mood, trait-related impairments may be a consequence of accumulating burden of illness on the brain. Previous research has similarly found that some cognitive impairments appear to be traits of the disorder: executive dysfunction may be a trait of MDD that is predicted by longer disease duration, whereas slower PS may be dependent on current state of the disorder, as is predicted by depression severity (Hu et al., 2022; J. Liu et al., 2019). Nilsson et al. (2016) found a relationship between attention and illness duration, but not severity, in MDD, suggesting that attention may be a

trait of the disorder and not state specific. However, both studies only analysed a depressed group, so more research is needed to verify this. In our results, we found that PS and EF contributed to memory impairments in euthymic BD, suggesting these may be trait-features of BD. In our sample of unmedicated depressed MDD patients, once SA was accounted for, PS and EF did not add anything of significance in explaining the memory impairment. In BD depression, SA was also related to memory. SA may therefore be a state-related feature of mood disorders, but further research is needed to confirm this.

6.2.4. Brain-cognition associations

Studies of the neural correlates of mood disorders generally quantify differences in brain morphology between patient and control groups and demonstrate structural abnormalities in patients (Hibar et al., 2018; Karantonis, Carruthers, et al., 2021; Z. Zhu et al., 2022). However, there is currently limited understanding of the functional implications of these structural abnormalities. A recent review suggested that cortical thickness appears to be related to cognitive function in BD, where those who are more cognitively impaired show a greater magnitude of brain structural abnormalities, however, there was not enough research to provide strong evidence for this (Karantonis et al., 2023). Ours was one of few studies to assess multivariate brain-cognition associations in BD and suggested that more abnormal structure in several regions was related to greater cognitive dysfunction. However, whether these brain-cognition associations reflect causal relationships is unclear: it may be the case the abnormal brain functioning leads to poorer cognitive function or vice-versa. Alternatively, both may be a consequence of disease burden. More research is needed to establish the underlying mechanisms driving brain-cognition associations in mood disorders.

We found that several brain regions may be related to core cognitive impairment in BD. This invites research from a whole-brain perspective that can assess the role of entire brain networks in cognitive functioning (Deng et al., 2019; McPhilemy, Nabulsi, Kilmartin, O'Hora, et al., 2020; van den Heuvel & Sporns, 2019). PS appears to depend on effective communication between brain regions (Ajilore et al., 2015; King & Anderson, 2018) and abnormal network properties have been linked to memory impairments and depression symptoms (G. H. Kim et al., 2019). Compared to healthy controls, people with BD show abnormalities in resting-state network characteristics (e.g., degree centrality) in neural emotion regulation circuit areas and in PS may be associated with degree centrality in the

superior temporal and inferior temporal gyri (Deng et al., 2019). Network analysis has also been applied to structural data: a study that assessed the network properties of the cortical thickness of regions in MDD showed that nodes (i.e., regions) that are associated with the DMN, salience network, and executive network, were disrupted in patients compared to controls (T. Wang et al., 2016). Further investigations of network properties may shed more light on the neural correlates of core cognitive dysfunction in mood disorder groups. Studies could use network analysis to calculate network properties for each region, based on either structural or functional data, and test the association of these network metrics with cognitive impairment using multivariate statistical models. Since network metrics offer a way of conceptualising brain structure from a whole-brain perspective (i.e., the relationship of one region to other regions), this would complement other research that considers brain data for each region separately.

6.3. Critical appraisal, limitations, and considerations

6.3.1. Strengths

Chapter 4 benefitted from the use of statistical methods to fractionate out cognitive components from raw neuropsychological data. For example, we removed the attentional/speeded scores from tests of EF where possible to better reflect effortful processing. We also used ex-Gaussian distributional parameters to measure SA, which deconstruct RT data into several components. These metrics have advantages in their ability to separate general PS components from longer responses thought to measure lapses in attention (Schmiedek et al., 2007; Whelan, 2008). Only one previous study to our knowledge has utilised these metrics in mood disorder groups (Gallagher, Nilsson, et al., 2015). Using these metrics, we were able to highlight SA as having a potential primary role in depressed states. SA had a large variance in the patient groups, suggesting it may be particularly sensitive to cognitive dysfunction in some mood disorder patients.

Our research also benefitted from the use of methods that investigated relationships between cognitive functions. Much of the research to date does not account for these interrelationships and uses univariate analysis which requires corrections for multiple comparisons (Abé et al., 2018; Hu et al., 2022; Krabbendam et al., 2000). We demonstrated the utility of multivariate methods to account for relationships between cognitive functions, which arguably better reflects complex, multicomponent data (Sherry & Henson, 2005).

6.3.2. Limitations

6.3.2.1. Heterogeneity in cognitive dysfunction

This research has some important limitations to consider when interpreting the results. One such limitation concerns the heterogeneity in cognitive impairment in patient groups, as there is considerable variation within and between cognitive functions in mood disorder groups (Cullen et al., 2016; Parkinson et al., 2020). While a group of patients may perform worse than controls on average, there exists some patients whose cognitive performance appears comparable to controls (Cotrena et al., 2017; Cullen et al., 2016; Douglas et al., 2018; Roux et al., 2019). Some studies suggested that only around 30-40% of people with mood disorders display cognitive impairment (Douglas et al., 2018; Iverson et al., 2011). It is possible that patients who appear cognitively intact had higher premorbid cognitive functioning and may experience cognitive decline at disease onset that is not detected in cross-sectional analysis, or may be countering cognitive impairments with a higher cognitive reserve (Gruber et al., 2022). However, the presence of some cognitively normal patients prevents us from concluding that cognitive dysfunction is characteristic of all mood disorder patients.

The finding that not all patients display cognitive impairment has led to investigations of the presence of distinct subgroups of patients with different cognitive profiles (Burdick et al., 2014; Wu et al., 2017). Studies using cluster analysis have grouped patients into three cognitive subtypes: those with global impairment, those with selective impairments in specific cognitive functions, and those who appear cognitively intact (J. Jensen et al., 2016; Karantonis et al., 2020; Lima et al., 2019; Miskowiak et al., 2023; Russo et al., 2017). The globally impaired subgroup tend to show poorer general functioning than other subgroups (J. Jensen et al., 2016; Miskowiak et al., 2023) and distinct cognitive profiles have been linked to differential impairments in affective cognition (Kjærstad et al., 2021). Other studies report four subgroups: one relatively intact subgroup, one globally impaired subgroup, and two selectively impaired subgroups (Lewandowski et al., 2018). The number of selectively impaired subgroups, and the nature of this selective impairment, is therefore not completely clear: some studies found a subgroup with selective impairments in working memory and PS (Miskowiak et al., 2023), whereas others found selective impairments in PS, attention, verbal learning and social cognition (Burdick et al., 2014). Further, while some studies suggest subgroups are best discriminated by PS (J. Jensen et al., 2016; Lima et al., 2019), others suggest

that subgroups cannot be differentiated by PS (Karantonis et al., 2020). Therefore, there may exist subgroups of patients with distinct cognitive profiles, however, findings on the exact nature of these profiles are mixed and further research is required.

Research should further investigate whether cognitive subgroups are differentiated by unique cognitive profiles, or whether cognitive subgroups represent an underlying continuum of cognitive dysfunction, where a similar profile of deficits exist over varying degrees of impairment (Gallagher, 2021; Lima et al., 2019; Van Rheenen et al., 2017). In the latter case, specific impairments found in the selectively impaired subgroup(s) may reflect the sensitivity of particular neuropsychological tests to cognitive impairment. Our results from Chapter 3 suggested that the magnitude of impairment in PS and SA varied depending on the outcome measure used, thus future research should carefully consider the neuropsychological methodology when investigating cognitive subgroups in mood disorders and interpreting apparent selective deficits. The existence of distinct cognitive profiles within patient groups is important to investigate as the presence of subgroups would indicate that cognitive remediation therapy should be targeted to specific cognitive functions in selectively impaired patients.

There may be other factors that can partly explain the wide variation in cognitive performance in mood disorder groups, including the clinical heterogeneity in the disorders. This limitation was accounted for in Chapter 5, where we assessed brain-cognition associations by treating abnormal cognitive functioning and abnormal brain morphology as continuous variables. However, the potential influence of cognitive heterogeneity should be considered when interpreting the results of Chapter 3 and Chapter 4, where we focussed on group means and conceptualised impairments as a significant difference from the control mean. Some potential sources of cognitive heterogeneity are discussed in the following section (6.3.2.2).

6.3.2.2. Clinical heterogeneity

The heterogeneity of cognitive performance in mood disorder groups may be driven by confounding clinical characteristics of the patients. Both MDD and BD are highly heterogeneous across individuals in terms of symptoms, severity of symptoms, and course trajectories (Wardenaar & de Jonge, 2013). The existence of clinical subtypes within each disorder further complicates research in cognitive functioning in these groups. Some research suggests that BD-I is associated with more severe cognitive impairment than BD-II, including

on tests of PS, verbal memory, EF, and global cognition (Bora, 2018; Cotrena et al., 2016); other research found no difference in cognitive impairment (Tsitsipa & Fountoulakis, 2015). Alternatively, cognitive subtypes may be better explained by a history of psychosis, which is linked to poorer cognitive functioning than BD patients without a history of psychosis (Bora, 2018). Other clinical characteristics appear to contribute to the heterogeneous findings in the literature, for example, more severe cognitive dysfunction has been associated with a higher number of mood episodes, higher sub-syndromal mood symptoms, younger age of illness onset, longer illness duration, and more hospitalizations (Bora, 2018; Cardoso et al., 2015; Goswami et al., 2006; MacQueen & Memedovich, 2017; Porter et al., 2015). Other studies found that illness severity, but not duration was associated with cognitive dysfunction in MDD (Hu et al., 2022). There are high rates of childhood trauma, such as emotional neglect or abuse, in people with mood disorders, and trauma has been related to cognitive impairment in mood disorders and may therefore play a role in cognitive heterogeneity (Bücker et al., 2013; Jørgensen et al., 2023; Mandelli et al., 2015; Watson et al., 2014). Sleep also appears to be related to neuropsychological functioning in mood disorders (Bradley et al., 2019; Pearson et al., 2021; M. Wang et al., 2022). Controlling for comorbidity has also been advised (Miskowiak et al., 2018); we set strict exclusion criteria for the systematic review and only included patients without comorbidities. However, many factors that we did not account for here may contribute to cognitive impairment and complicate investigations into the precise cognitive profile of mood disorders.

Clinical characteristics may also explain heterogeneity in brain structural abnormalities in mood disorders. Longer duration of illness, age of onset, number of episodes, and history of psychosis have all been related to decreased cortical thickness, even after accounting for age (Foland-Ross et al., 2011; Hanford et al., 2016; Hibar et al., 2018; Oertel-Knöchel et al., 2015). Mood state may also be linked to brain structure, with differences in cortical thickness of some areas between euthymic and depressed BD (Kang et al., 2022). Some studies have suggested no difference in cortical thickness between BD-I and BD-II (Kang et al., 2022), however, one study found that BD-I and BD-II had distinct relationships between EF and cortical thickness (Abé et al., 2018). Other factors such as the presence of childhood trauma may also drive heterogeneity in brain structure, as childhood trauma has been associated with brain structural abnormalities in mood disorders (Jørgensen et al., 2023). We attempted to control for known demographic confounds in our other analyses, such as age and sex (Z. Zhu et al.,

2022), however we did not control for clinical characteristics, so these factors may have affected our results. Other studies have found brain-cognition associations even after controlling for clinical characteristics (Macoveanu et al., 2021); it is still possible, then, that our findings exist independent of clinical confounds. The nature of the relationships between cognitive function, brain morphology and clinical characteristics, how much these overlap, or whether they reflect the same underlying pathology, is yet to be uncovered. The potential effects of demographics and clinical differences in our patient samples should not be discounted.

6.3.2.3. Medication

Antidepressant medication appears to have an effect on improving cognitive functions (Miskowiak et al., 2018; Porter et al., 2015). Since we did not control for medication in our systematic review, this could have led to heterogeneity in the results. However, some studies did not find an effect of medication on cognitive performance (Goswami et al., 2009), therefore more research is needed to confirm whether medications may confound results of neuropsychological assessments such as ours.

Medications are also associated with cortical thickness in BD (Hibar et al., 2016, 2018). Medication may a notable confound in our brain-cognition analysis: approximately half of the BD patients were taking lithium and the other half were taking maintenance treatments but were naïve to lithium. Many studies have suggested that, compared to patients not taking lithium, patients taking lithium have better EF, greater cortical thickness and surface area in several areas, and greater subcortical volumes of the hippocampus, putamen, thalamus (Hibar et al., 2016, 2018; Ortiz et al., 2021). Lithium may therefore have a neuroprotective effect and since half our sample were taking lithium, this may be a factor in our study. Specifically, it may have been the case that patients taking lithium had greater cortical thickness than those not taking lithium, which may have brought the patient group mean scores closer to that of the control group. However, we did not explore this potential confound for several reasons: first, previous analysis of a sample which overlaps with ours suggested that there was no difference in T1-relaxometry between the lithium subgroup and the non-lithium subgroup after correction for multiple comparisons (Necus et al., 2021), suggesting our medication subgroups may not show different brain morphology. Secondly, our sample size was relatively small (*n*=56 patients), so further separating this into two subgroups would have further reduced

statistical power. Finally, while the presence of patients with lithium may have influenced the average cortical thickness of the whole patient group, our analysis was not based on group means, and instead we assessed cortical thickness and cognitive impairments as continuous variables. Since our analysis was based on correlations between cognitive impairment and cortical thickness, effects of lithium on mitigating both cortical thinning and cognitive impairment in our sample would have been in line with this correlation, i.e., these two variables would still be associated in lithium patients, regardless of a potentially smaller magnitude of abnormality. Therefore, potential differences in cortical thickness between lithium and non-lithium patients may not have been detrimental to the results. However, it is still possible that the nature of the continuous brain-cognition associations may change in the presence of lithium; more research with larger samples is required to test the effect of medication on brain-cognition associations.

6.4. Open questions and future directions

6.4.1. Neuropsychological methodology

As mentioned above (section 6.2.1), our research highlights the importance of careful consideration of the neuropsychological methodology used to measure cognitive functions in mood disorders. Efforts should be made to mitigate the task impurity problem by fractionating out, as much as possible, other cognitive functions that tests may capture, for example, removing speeded and attentional components from tests of EF. Our review suggested that the most common measures of PS are graphomotor tests such as DSST or TMT-A; however, since these tests may be limited by the task-impurity problem, PS may be better conceptualised by using RT-based methods. For example, the Hick-Hyman law states that RT increases as a (logarithmic) function of the number of response alternatives (Hyman, 1953; Proctor & Schneider, 2017). Choice-RT tasks with a varying number of response alternatives could therefore be utilised to measure baseline PS (i.e., trials with one response alternative; simple-RT), as well as measuring the slope of the increase in RT as task complexity (response alternatives) increases. This would reflect the speed at which individuals can processes increasing levels of information. SA may be better represented using sophisticated models of RT from CPTs, such as ex-Gaussian modelling. Future research should focus on SA in depression given that this may be a primary impairment in depressed states that may be related to mind-wandering.

6.4.2. Other structural metrics

In this study, we chose to measure cortical thickness as an index of brain structure. Cortical thickness was chosen for several reasons: first, due to its known association with cognitive impairment in mood disorder groups (Karantonis et al., 2023; Macoveanu et al., 2021). Secondly, while other morphological metrics such as cortical volume and fractal dimensionality may have stronger correlations with age than cortical thickness, the effect of age was not a variable of interest in our study, and we controlled for age in our analysis. Finally, cortical thickness is a commonly used structural brain metric in the literature, whereas less traditional morphological metrics such as fractal dimensionality are rarely used, therefore using cortical thickness allows us to compare our results to previous research. Since there is not yet a consensus on which brain regions may be related to cognitive impairment in terms of cortical thickness, the ability to compare our results, which used a novel statistical method (CCA), to previous research was important in interpreting our results. However, future research may benefit from using other morphological metrics in multivariate brain-cognition analyses to compare those results to ours.

While recent research suggests that cortical thickness may be more strongly related to cognitive impairment than cortical volume (Karantonis et al., 2023; Macoveanu et al., 2021), other morphological metrics, such as cortical volume and surface area, should also be explored. Gyrification may be another useful measure, as reduced global cortical folding has been found in depression (Penttilä et al., 2009) and reduced gyrification in the PFC was associated with poorer EF in BD (A. McIntosh et al., 2009). Fractal dimensionality, which represents a mathematical measure of the complexity of a structure, has been shown to have stronger age correlations than cortical thickness and gyrification (Madan & Kensinger, 2018). Differences in fractal dimensionality have been shown in people with anorexia compared to HCs (Collantoni et al., 2019). This may therefore be useful tool to for investigating brain structure in people with mood disorders.

Most research uses single metrics such as cortical thickness, cortical volume, and surface area to compare brain structure of clinical groups and controls, however, these metrics are known to covary, which perhaps limits the interpretability of each metric when used alone (Y. Wang et al., 2020). The independent components of cortical brain morphology developed by Wang et al. may represent a more valid measure of structural morphology that accounts for covariance between individual features. The components appear to be sensitive to structural

differences in other clinical groups, such as Temporal Lobe Epilepsy (Wang *et al.*, 2020). Future research should consider the use of these independent components to compare morphology in BD and MDD. Such techniques would add to the more traditional imaging metrics to build a more in-depth picture of brain structure and functioning in mood disorders.

Future studies should also consider the role of structural connectivity in core cognitive impairment in BD. People with mood disorders show reduced white matter integrity (Kempton, Geddes, Ettinger, Williams, & Grasby, 2008; Lloyd et al., 2009; Wise et al., 2016), which appears to be related to cognitive dysfunction (Kieseppä et al., 2014; Macoveanu et al., 2021; Rizk et al., 2017), including attention and SA (Masuda et al., 2020; Poletti et al., 2015). White matter integrity can predict sub-types of BD with high accuracy (Wu et al., 2017) and is associated with clinical characteristics such as late onset, short disease duration, and medication (Favre et al., 2019), and may therefore be related to cognitive heterogeneity. Reduced connectivity in BD has also been found using structural covariance methods (i.e., cortical inter-regional correlations; Kuang et al., 2022) and may therefore be an important feature of mood disorders relevant to cognitive impairment.

6.4.3. DMN and functional imaging

Given that SA is closely associated with mind-wandering (Esterman & Rothlein, 2019; Fortenbaugh et al., 2017), our finding that SA may be a primary deficit in depression is consistent with the link between depression and mind-wandering and rumination (Rosenbaum et al., 2017; Seli et al., 2019). Our brain-cognition analysis suggested that structural abnormalities of DMN-related areas may be related to core cognition in BD, which is in line with the association between SA and the DMN in controls (Esterman et al., 2013) and well as the link between DMN and rumination in depression (Dutta et al., 2019; Rosenbaum et al., 2017). Together, our results provide tentative support for the role of the DMN in depression, its link to SA and the effect of SA on wider functioning. This calls for further research into these relationships using ex-Gaussian measures of SA to fractionate out PS.

The potential link between the DMN and cognitive functioning in mood disorders has been studied in previous research of brain function in patients: people with BD and MDD show abnormal functional connectivity of the DMN, including failure to suppress the DMN during tasks and abnormal variability in activity of the DMN, which may reflect a state of rigid connectivity (Miskowiak & Petersen, 2019; Rashid et al., 2014; Zarp Petersen et al., 2022).

Similarly, BD have shown less variable temporal connectivity strength between the medial PFC and other regions of the DMN, specifically the PCC, compared to HCs (Nguyen et al., 2017). Further, a higher degree of this abnormality (i.e., less variable connectivity) was associated with slower PS and poorer EF in BD. Nguyen et al. (2017) suggested that lower neural variability may reflect an inflexibility to switch between networks (i.e., lower inter-network flexibility). This compliments the findings that in HCs, activity in regions thought to be inter- or intra-network hubs is related to cognitive performance in general (Burzynska et al., 2015), and that less flexible activity between regions is associated with slower processing (King & Anderson, 2018). This finding may be of particular importance in implicating regions such as the PCC, which are thought to be 'hubs' central to several networks in the brain with a role in internally-directed cognition, attention allocation, and coordination of various networks (Leech, Kamourieh, Beckmann, & Sharp, 2011; Leech & Sharp, 2014). Abnormal connectivity of the PCC may therefore reflect reduced ability to allocate attention and react quickly, this may contribute to poor attention and PS in BD. While SA has been related to altered resting-state functional activity between regions in BD, regardless of mood state (Yu et al., 2021), the specific link between SA and DMN activity has not yet been fully explored.

Another region which appears to have a role in cognitive impairment in BD is the dorsal PFC, which is involved in the cognitive control network (CCN) (Cole & Schneider, 2007). BD patients show hypoactivity in this region, which has been associated with impairments in working memory and memory (Frangou et al., 2008; Macoveanu, Mariegaard, et al., 2023; Macoveanu, Petersen, et al., 2023). In line with this, recent research using structural metrics found that cognitively impaired BD patients show abnormal dorsal PFC cortical thickness compared to cognitively normal patients and healthy controls (Macoveanu et al., 2021). Our analysis in Chapter 5 did not detect an association between abnormal dorsal PFC thickness and cognitive impairment, perhaps because our neuropsychological battery did not include tests of working memory or memory. A recent meta-analysis found that the most consistent changes in neural circuitry in BD patients who are cognitively impaired were aberrant activity in dorsomedial and dorsolateral PFC regions and changes in the DMN (Miskowiak & Petersen, 2019). The dorsal PFC and the DMN may therefore both have a role in cognitive impairment in BD, where there is *hypo*-activity in the CNN and *hyper*-activity in the DMN (Zarp Petersen et al., 2022).

Future studies should therefore investigate the association between core cognitive impairment and functional brain activity in mood disorders, focussing on the role of the PCC

in the DMN, and the dorsal PFC in the CNN, and how these relate to poor cognitive functioning, particularly SA. Such studies may benefit from the use of network models to assess brain-cognition associations from a whole-brain perspective (van den Heuvel & Sporns, 2019); previous network analysis has demonstrated that BD show differences in network metrics compared to controls (Deng et al., 2019; McPhilemy, Nabulsi, Kilmartin, O'Hora, et al., 2020).

6.4.4. Other mood disorders

In this study we assessed brain-cognition associations in BD only. It may be useful to compare such results across groups, for example, comparing BD with HCs. Whether BD is associated with unique brain-cognition associations or whether these associations are just an extension of HC models, i.e., greater structural integrity equals better cognitive functioning, is unclear. Some studies suggest that brain-cognition associations in BD are similar to HCs, suggesting BD are on the 'poorer' end of the spectrum than HCs (Fears et al., 2015). Future research should also investigate brain-cognition associations in MDD. Previous research suggests that cortical thickness may be decreased in BD compared to MDD (Lan et al., 2014), so brain-cognition associations may differ between the diagnostic groups. Canonical discriminant analysis may be a useful tool for this, which has previously been used to compare brain-cognition associations between groups across the affective psychosis continuum (Rodrigue et al., 2018).

6.4.5. Longitudinal prospective studies

Much of the research in the literature, and in this thesis, is cross-sectional in nature and therefore causation cannot be inferred from the results. However, when considering a hierarchy of cognitive functions, or how cognitive functions relate to brain structure, it would be useful to examine these variables longitudinally. For example, in the case of a hierarchy of cognitive dysfunction it is important to determine whether the effect of primary impairment(s) on wider cognitive function occurs slowly over time, or whether it happens simultaneously. Here, longitudinal prospective cohorts would be useful: future studies should recruit first-episode patients and assess the relationship between their cognitive functions over time to test whether a hierarchy of dysfunction is present at first-episode, or if it develops over time. It would also be beneficial to track those that are at a higher risk of developing mood disorder, such as first-degree relatives of patients, to assess the nature of cognitive functions in high-risk groups, and then after diagnosis in those who go on to develop BD or MDD. Prospective studies have tracked cognitive impairment over time (Kjaerstad et al., 2023;

Sparding et al., 2021), however, future studies should also track the *interaction* of cognitive functions over time in order to assess the possible trajectory of a hierarchy of cognitive dysfunction. Detecting primary impairments that occur early in the disease process and go on to affect wider functioning could allow cognitive remediation therapies to target specific functions at an early stage before the disease progresses and potentially prevent further cognitive decline.

Similarly, for brain-cognition associations, prospective longitudinal cohorts, including high-risk groups, would be useful for identifying abnormal brain-cognition associations that are present before diagnosis or at first episode, which may reflect traits of the disorders, or early biomarkers. Further, the trajectory of brain-cognition associations could be tracked to investigate how these interact over time. For example, it could be that cognitive deficits lead to abnormal brain structure due to patients recruiting different brain regions in a form of 'cognitive scaffolding', where performance on some impaired functions are supported by other processes (Gallagher, Gray, et al., 2015). Alternatively, cognitive functions may follow from brain abnormalities caused by neurobiological pathology such as an increased inflammatory state (Van Rheenen et al., 2020). It may also be possible that cognitive deficits and brain structural abnormalities are both independently affected by disease burden. The literature of the trajectory of cognitive impairment in mood disorders, and whether it is neuroprogressive, is mixed (Van Rheenen et al., 2020), with some studies showing that cognitive deficits are stable over time in patients and unaffected first-degree relatives (Kjaerstad et al., 2023; Sparding et al., 2021), and others suggesting there is progression of cognitive dysfunction in a subgroup of patients with more (hypo)manic episodes (Sánchez-Morla et al., 2019). Future research should therefore implement prospective longitudinal studies to track cognitive hierarchy and brain-cognition associations over time.

6.5. Conclusion

This thesis suggests that, compared to healthy controls, people with BD and MDD show impairments in PS and SA that are present in euthymia/remission. Impairments in PS, SA, and EF may be primary impairments in mood disorders that could lead to secondary cognitive impairment. However, the specific cognitive interrelationships that are important in each disorder are not yet clear, and more research is needed to decipher the precise cognitive profile of BD and MDD and compare this across different mood states. Core impairments also

appeared to be related to abnormal brain structure in BD and our study highlighted the utility of CCA to test brain-cognition associations in psychiatric samples. Future work should use multivariate methods to replicate our findings and probe the specific brain-cognition associations in BD and MDD.

Uncovering the cognitive profile of mood disorders could inform clinical practice by targeting primary cognitive impairments associated with BD and MDD in cognitive interventions or with medications such as vortioxetine and pramipexole that have been found to improve processing speed in MDD (Baune et al., 2018; Rosenblat et al., 2016). Similarly, revealing how cognitive impairments are related to brain structure and function in mood disorders could help to disentangle cognitive and neural features that are related to mood states and features that are traits of each disorder. While research into brain-cognition associations in mood disorders is at an early stage that cannot yet inform clinical practice, research should continue to tease apart structural and functional brain abnormalities that are linked to cognitive impairment. For example, studies should use hierarchical regression in larger samples to identify primary cognitive impairments and then use multivariate methods in large datasets to test which brain regions are associated with these primary impairments in terms of both brain structure and function. The output of these studies could indicate primary impairments in BD and MDD that could later be targeted in randomised controlled trials of cognitive remediation therapy, and understanding the neural correlates of primary impairments would inform our understanding of the mechanisms of cognitive remediation. Finally, uncovering reliable brain-cognition associations could provide useful biomarkers for BD and MDD which could, with more research, aid clinical diagnosis.

APPENDIX A

a) ELIGIBILITY CRITERIA AND EXCLUSION REASONS

	Include	Exclude	Reason for exclusion (code)
Type of publication	Original articles	Reviews Systematic Reviews Meta-analyses Theses/Dissertations Abstracts Conference proceedings Study protocols Posters Case studies Commentaries/Opinion articles/editorial Guidelines, other reports Book or book chapter Other grey literature	Review Systematic Review Meta-analysis Thesis Abstract Conference Protocol Poster Case study Opinion Guideline Book/chapter Other (not original article)
	Published	Pre-print Unpublished	Unpublished
	English language (full text)	Any other language with no English translation of the full text	Not English
	Published from 1994 onwards	Published before 1994	Pre-1994
	Published in a peer-reviewed journal	Not in a peer-reviewed journal	Not peer-reviewed
Design	Investigated neuropsychological function as a primary aim	Does not measure neuropsychological function as primary aim	No neuropsych Not primary aim
	Not primarily an imaging study or eye tracking study	Primarily imaging studies (MRI, fMRI, EEG, MEG, PET, TMS) Primarily eye tracking studies	Imaging Eye tracking
	Cross-sectional studies or baseline results from longitudinal, cohort studies or trials	Other, e.g., case-study	Not cross-sectional data
Methods	Included at least one objective test of processing speed and/or at least one test of sustained attention	Does not contain any measure of processing speed or sustained attention (even if there are tests of other cognitive domains) Subjective cognitive tests (e.g., self-report questionnaires)	No tests of PS or SA Subjective cognitive test
	The test should yield a numeric score, or a rating (e.g., pass/fail,	Non-numeric score	Non-numeric test score

	or poor/fair/good, or impaired/unimpaired)		
Sample	Humans	Animal models Computer models	Not humans
	Adults (18-65)	Children or youths (<18, paediatric, adolescent)	Children
		Older adults (65+, Late-Life Depression, elderly)	Older adults
	Contains a sample of people with a primary diagnosis of Major Depressive Disorder (MDD; depression) and/or a sample of people with a primary diagnosis of bipolar disorder (BD)	HCs only (even if they report a range of mood scale scores)	No MDD or BD (HCs only)
		First degree relatives of people with mood disorders	No MDD or BD (FDRs only
		People with a diagnosis of any other psychiatric disorder (e.g., schizophrenia, schizoaffective disorder, seasonal affective disorder, post-natal depression, premenstrual dysphoric disorder. depression related to medical illness, depression induced by substance use or medication)	No MDD or BD (other psych disorder)
		Any other clinical group	No MDD or BD (other clinical group)
		Not primary diagnosis	Not primary diagnosis
	MDD or BD diagnosed using a recognised criterion-based diagnostic system (DSM or ICD) (Ok if diagnosed in the past; mood scales not sufficient)	Unclear, insufficient, or non-clinical diagnostic criteria	Diagnosis tool
	Any mood state (euthymic, depressed, manic)	-	-
	No severe psychiatric comorbidity or neurological disorder or intellectual disability	Severe psychiatric comorbidity (e.g., autism, schizophrenia) or neurological disorder (e.g., dementia, epilepsy, stroke) or intellectual disability	Comorbidity
	Contains a HC group	No HC group	No HC
Data	Reported uncorrected mean test score and standard deviation (SD) for both patient and control groups for each neuropsychological test	Means and SDs not available for patient and control groups	Data missing
		Domain-level composite data only	Composite data only
	Original sample (i.e., data from sample not published elsewhere)	Redundancy with other retained papers	Redundant

Table 18: Eligibility criteria and exclusion reasons.

b) FINAL SEARCH STRATEGY

#	Search
1	(Neurocog* or neuropsychol* or cogniti* or attention* or 'sustained attention' or vigilance or 'mental speed' or 'processing speed' or 'speed of processing' or 'speed of information processing' or psychomotor).kf,kw,ti.
2	(bipolar or manic or mania or manic-depress* or BD or depress* or MDD or 'mood disorder*' or 'affective disorder*).kf,kw,ti.
3	1 and 2
4	limit 3 to English language and yr="1994 -Current"
5	('attention deficit hyper*' or ADHD or 'ADD' or 'attention deficit disorder*').kf,kw,ti.
6	(therapy or CBT or 'cognitive-behavio*').kf,kw,ti.
7	'dement*'.kf,kw,ti.
8	4 not 5 not 6 not 7
9	limit 8 to journal article [Limit not valid in Embase, Journals@Ovid; records were retained]
10	limit 9 to peer reviewed journal [Limit not valid in Ovid MEDLINE(R), Embase, Journals@Ovid; records were retained]
11	limit 10 to journal [Limit not valid in Ovid MEDLINE(R), Journals@Ovid; records were retained]

Table 19: Final search strategy as implemented in Ovid, including database-specific restrictions (9-13).

c) DATA EXTRACTION FORM

Type of data	Data collected	Class of data
Citation details	First author	Open text
	Year of publication	Year
	Paper title	Open text
	Journal	Open text
	Corresponding author email address	Open text
Paper details	Country of publication	Open text
	Study design	Open text
	Setting (general community, clinical out-patient, hospital in-patient)	Categorical
Methods	Method of recruitment for clinical group and comparison groups	Open text
	Diagnostic criteria (and tool used) for MDD/BD	Open text
	Mood state criteria	Open text
	Clinical tools (e.g., to measure illness severity)	Open text
	Eligibility criteria applied for clinical group and comparison group	Open text
Clinical characteristics	Clinical group (MDD, BD, or both)	Categorical
	Clinical subtype (e.g., BD-I or BD-II)	Open text
	Mood state (at time of testing)	Categorical
	Medication details	Open text
	Duration of illness	Numeric
	Age of illness onset	Numeric
	Mania scale score (mean, SD)	Numeric
	Depression scale score (mean, SD)	Numeric
Sample demographics (for each group)	Comparison group matching	Open text
	N of each group	Numeric
	Age (mean and SD)	Numeric
	Gender (% female)	Numeric
	Years of education	Numeric
	Ethnicity	Open text
Neuropsychological outcome variable data	Neuropsychological domains measured (as reported by authors)	Numeric
	Neuropsychological tests used	Open text
	Source reference for test	Open text
	Outcome variables used for each test	Open text
	Characteristics of test (e.g., length, number of trials, blocks)	Open text

	Cut-offs used for impaired/not impaired	Open text
	Mean and SD of each test score	Numeric
	Did the authors find a significant group difference	Categorical
Risk of bias information	Power analysis	Open text
	Counterbalancing	Open text
	Missing data	Open text
	Statistical analysis	Open text
	Conflict of interest	Open text
	Funding Sources	Open text
Other	Any other notes	Open text

Table 20: Data extraction form headings.

d) PAPER CHARACTERISTICS

Study	Mood disorder	Mood state	Diagnostic criteria	Other clinical characteristics	Setting	Patient N	Patient age, Mean (SD)	Patient sex, % female	Control N	Test(s) of PS	Test(s) of SA
Arslan 2014	BD	euthymic	DSM-IV	n/a	outpatients	30	33.37 (10.09)	67%	32	TMT-A	n/a
Bakusic 2021	MDD	depressed	DSM-IV	n/a	inpatients	79	45.00 (11.90)	58%	57	TMT-A	n/a
Behnken 2013	MDD	remitted	DSM-IV	First episode. Non-psychotic.	outpatients	20	37.27 (10.70)	65%	20	TMT-A	n/a
Bhardwaj 2010	MDD	remitted	DSM-IV	n/a	outpatients	20	34.30 (8.20)	10%	20	DSST	n/a
Booij 2006	MDD	remitted	DSM-IV	Non-psychotic. Unmedicated.	outpatients	23	29.96 (9.70)	91%	20	Stroop task, Choice-RT	n/a
Bradley 2019	BD	various	DSM-IV	n/a	outpatients	46	47.50 (12.55)	50%	36	PVT, DSST, TMT-A	ANT
Braund 2020	MDD	depressed	DSM-IV	Non-psychotic.	outpatients	766	37.73 (12.31)	57%	336	Choice-RT	CPT
Burdick 2009	BD	depressed	DSM-IV	Non-psychotic.	outpatients	24	39.96 (9.50)	38%	24	Simple-RT, Choice-RT	n/a
Burdick 2014	BD	euthymic	DSM-IV	Some with psychotic features.	outpatients	136	40.80 (10.60)	50%	148	MARTRICS PS (BACS, TMT-A)	CPT-IP (MCCB)
Castaneda 2008	MDD	not stated	DSM-IV	Non-psychotic.	community	46	28.50 (3.80)	74%	70	DSST (WAIS-R), TMT	n/a
Chang 2012	BD	Euthymic	DSM-IV-TR	Psychotic symptoms.	inpatients and outpatients	34	BD-I=32.64 (11.18), BD-II=28.00 (4.89)	BD-I=36, BD-II=61%	30	DSST (WAIS-III), TMT-A	n/a
Cheung 2013	BD	euthymic	ICD-10	Some with psychosis.	outpatients	52	38.57 (10.70)	63%	52	CNSVS: PS composite, DSST	CNSVS Complex Attention
Daniel 2013	BD and MDD	euthymic/ remitted	DSM-IV-TR	Stable medication. Non-psychotic.	outpatients	50	50.60 (8.28)	64%	29	DSST (WAIS-R)	n/a
den Hartog 2003	MDD	depressed	DSM-IV	Moderate to severe MDD. Non psychotic. Unmedicated.	outpatients	30	41.60 (12.40)	47%	38	Stroop task, TMT-A (concept	n/a

										shifting task), Verbal fluency	
Donohoe 2012	BD	not stated	DSM-IV	n/a	not stated	110	44.82 (10.50)	46%	163	n/a	CPT
Doose-Grünefeld 2015	MDD	depressed	DSM-IV	n/a	inpatients and outpatients	41	36.49 (10.87)	46%	41	TMT-A	n/a
Douglas 2011	BD and MDD	depressed	DSM-IV	Some with psychotic features, some melancholic.	inpatients	*n*=60 MDD, *n*=8 BD	MDD=38.90 (10.80), BD=41.80 (11.70)	MDD: 62%; BD: 63%	50	Simple-RT, Timed Chase Test, Stroop task, Verbal fluency	n/a
Duan 2021	MDD	depressed	DSM-IV	Unmedicated.	not stated	221	37.13 (10.81)	72%	499	Verbal fluency, DSST, TMT-A (Colour trial test [median and range scores])	CPT (not meta-analysed; only provided median and range scores)
Elshahawi 2011	BD	euthymic	ICD-10	50% single manic episode, 50% recurrent episodes. History of psychotic features.	outpatients	100	31.25 (8.05)	37%	50	TMT-A	n/a
Erol 2014	BD	euthymic	DSM-IV	n/a	outpatients	25	30.60 (6.40)	32%	25	TMT-A, Stroop task	n/a
Esan 2020	BD	euthymic	DSM-IV	n/a	outpatients	110	39.10 (11.10)	62%	100	DSST (SCIP processing speed), Verbal fluency (SCIP)	n/a
Fernández-Sevillano 2021	MDD	depressed	DSM-5	Non-psychotic.	inpatients	76	46.37 (11.21)	75%	20	PS composite: WAIS-IV Digit symbol and coding	n/a
Ferrier 1999	BD	euthymic	DSM-IV	Minimum 5-year history of illness.	not stated	41	44.72 (10.52)	66%	20	Digit symbol, TMT-A	n/a
Frajo-Apor 2020	BD	various	DSM-IV	Some with psychotic episode.	outpatients	54	45.90 (11.40)	93%	80	DSST	n/a

Frydecka 2014	BD	symptomatic (depressed, manic)	DSM-IV	Some psychotic.	outpatients	43	43.76 (4.67)	40%	18	n/a	CPT (AX)
Gallagher 2014	BD	depressed	DSM-IV	No current psychosis.	outpatients	53	47.00 (10.00)	38%	47	DSST, SCOLP, Verbal Fluency	CPT (Vigil)
Gallagher 2015	BD and MDD	depressed (MDD); depressed and euthymic (BD)	DSM-IV	n/a	outpatients	*n*=86 BD-e, *n*=33 BD-d, *n*=39 MDD	BD-e=44.00 (9.74), BD-d=47.00 (8.64), MDD=32.30 (10.11)	BD-e=52, BD-d=42, MDD=62	138	n/a	Vigil CPT
Gogos 2010	BD	euthymic	DSM-IV	n/a	outpatients	40	42.40 (11.65)	60%	43	DSST (RBANS coding)	n/a
Goltermann 2021	MDD	not stated	DSM-IV	n/a	not stated	547	38.35 (13.60)	57%	670	TMT-A, DSST	n/a
Gómez-Benito 2014	BD	euthymic	DSM-IV-TR	n/a	outpatients	76	40.30 (8.98)	74%	83	DSST (SCIP Processing Speed Test), TMT-A	n/a
Gorenstein 2006	MDD	not stated	DSM-IV	n/a	outpatients	56	40.70 (1.40)	75%	31	DSST, Simple-RT (Foundations I, PSS)	n/a
Gu 2016	MDD	depressed	DSM-IV	First episode.	inpatients and outpatients	100	28.03 (7.05)	47%	46	Verbal fluency, DSST (WAIS-RC), TMT-A, Stroop task	n/a
Gualtieri 2008	BD and MDD	not stated	DSM-IV-TR	Some not medicated.	not stated	455 (*n*=336 MDD, *n*=119 BD)	BD=37.20 (13.08), MDD=40.10 (11.35)	BD=62, MDD=65	336	PS composite (CNSVS; FTT and DSST)	SA Other (CNSVS Complex Attention: Stroop, CPT)
Halappa 2018	MDD	depressed	DSM-IV	Non-psychotic. Unmedicated.	outpatients	65	34.60 (8.85)	42%	19	TMT-A	n/a

Halvorsen 2012	MDD	depressed, remitted (separate)	DSM-IV	Non-psychotic.	outpatients	118 (*n*=37 depressed, *n*=81 recovered)	depressed=37.49 (11.98), recovered=37.42 (9.61)	depressed=73, recovered=88	50	Verbal fluency (D-KEFS), CalCAP Simple-RT and Choice-RT, DSST (WAIS-III), Stroop test (D-KEFS), TMT-A	n/a
Harmer 2002	BD	euthymic	DSM-IV	Diagnosed with BD for at >2 years.	inpatients	19	38.40 (11.33)	53%	19	n/a	CPT (non-working memory vigilance task)
Holmes 2008	BD	depressed	DSM-IV	n/a	not stated	65	38.24 (10.23)	65%	52	n/a	CPT (CANTAB RVIP)
Hou 2020	MDD	depressed	ICD-10	Episode lasting for >1 month.	outpatients	96	29.53 (10.92)	61%	97	THINC-it Spotter (Choice RT), Codebreaker (DSST)	n/a
Hsu 2015	MDD	not stated	DSM-IV	Non-psychotic. Drug-naïve.	outpatients	30	29.40 (6.47)	60%	30	n/a	CPT
Ji 2020	MDD	depressed	ICD-10	Drug-naïve. Starting a new antidepressant monotherapy.	outpatients	67	31.00 (9.84)	55%	56	DSST	n/a
Jiménez-López 2017	BD	euthymic	DSM-IV	50% history of psychosis, 50% no psychosis.	outpatients	100	41.80 (11.60)	50%	51	TMT-A, DSST (WAIS-III), Verbal Fluency	CPT (Degraded Stimulus)
Jin 2020	MDD	depressed	DSM-IV	First episode. Drug naïve.	not stated	100	27.79 (7.17)	62%	100	MCCB PS composite (TMT-A, DSST, verbal fluency)	CPT (MCCB attention/ vigilance)
Kim 2014	BD	euthymic	DSM-IV	n/a	inpatients and outpatients	28	36.68 (8.17)	46%	28	Verbal fluency, TMT-A	CPT (Degraded Stimulus)
Kim 2015	BD	euthymic	DSM-IV	History of psychotic symptoms; not currently psychotic.	outpatients	34	31.40 (7.50)	68%	34	Verbal fluency	CPT

Koopowitz 2021	MDD	not stated	DSM-IV	Pregnant women. Non-psychotic.	community	30	28.17 (6.24)	100%	86	DSST (oral symbol digit), Choice-RT (pattern comparison), TMT-A (colour trails test 1)	n/a
Lee 2017	BD	not stated	DSM-IV	First-diagnosed, medication naïve.	inpatients and outpatients	32	37.90 (10.80)	66%	30	BACS symbol coding (DSST), Verbal fluency	n/a
Leung 2016	BD	euthymic	DSM-IV	Some had history of psychosis.	outpatients	30	34.13 (12.48)	60%	30	n/a	CPT (SART)
Levada 2019	MDD	depressed	DSM-5	Unmedicated.	outpatients	119	39.10 (11.90)	68%	71	DSST	n/a
Lewandowski 2020	BD	not stated	DSM-IV	Psychosis.	outpatients	119	28.90 (9.50)	47%	87	MCCB: TMT-A, DSST, Verbal fluency	CPT (IP; MCCB)
Lewandowski 2016	BD	not stated	DSM-IV	Psychosis.	inpatients and outpatients	42	29.60 (8.40)	55%	29	MCCB: TMT-A, DSST, Verbal fluency	CPT (IP; MCCB)
Liang 2020	BD and MDD	depressed (MDD), euthymic (BD) (separated)	DSM-5	n/a	not stated	91 (*n*=43 BD, *n*=48 MDD)	BD=36.30 (11.90), MDD=39.70 (13.40)	BD=47, MDD=67	35	MCCB PS composite	CPT (IP; MCCB)
Lima 2019	BD	euthymic	DSM-5	Some had psychotic symptoms in first episode.	not stated	73	47.28 (12.15)	69%	57	Stroop task, TMT-A, Verbal fluency	CPT (IP; WASI-III)
Lin 2021	MDD	depressed	DSM-IV	94% unmedicated.	inpatients and outpatients	639	39.20 (10.40)		287	Verbal Fluency, DSST (BACS), TMT-A (Color trial test I), Stroop task	CPT (IP)
Liu 2019	BD and MDD	depressed	DSM-IV	Drug-naïve.	inpatients and outpatients	60 (*n*=30 BD,	BD=25.63 (6.65),	BD=63, MDD=67	30	DSST (WAIS-RC)	n/a

						n=30 MDD)	MDD=27.77 (7.15)				
Lu 2021	BD and MDD	depressed	DSM-IV	Some with psychotic symptoms.	not stated	90	BD median=28.5 (17,33), MDD median=21 (19,40)	BD=47, MDD=57	30	DSST (THINC-it), Choice-RT (THINC-it)	CPT
Lyness 1994	MDD	depressed	DSM-III-R	Medication-free for at >2 weeks.	outpatients	19	51.60 (6.70)	42%	16	TMT-A, DSST, Verbal fluency	n/a
Mahlberg 2008	BD	manic	DSM-IV	n/a	inpatients	30	46.40 (13.40)	53%	30	TMT-A	n/a
Marotta 2015	BD	euthymic	DSM-IV-TR	n/a	inpatients and outpatients	27	40.59 (13.11)	56%	27	n/a	SA other (ANTI-Vigilance)
McIntyre 2017	MDD	depressed	DSM-IV-TR	n/a	outpatients	100	40.68 (13.68)	51%	100	THINC-it Spotter (Choice-RT), Codebreaker (DSST)	n/a
Menkes 2019	BD	not stated	DSM-IV	Psychotic features.	not stated	112	32.30 (13.31)	51%	261	DSST (SCIP psychomotor speed)	n/a
Milas 2019	BD	not stated	ICD-10	At the initial stage or the disease.	not stated	18	38.80 (8.30)	100%	26	TMT-A	n/a
Miyata 2018	MDD	remitted	DSM-IV	n/a	outpatients	70	41.7.0 (7.20)	9%	67	TMT-A	CPT (IP)
Naim-Feil 2016	MDD	depressed	DSM-IV	Not psychotic. No response to >1 antidepressant.	not stated	21	44.00 (9.00)	52%	26	n/a	CPT (SART)
Nehra 2006	BD	euthymic	ICD-10	Some Psychotic.	inpatients and outpatients	46	33.52 (11.23)	33%	20	TMT-A, Verbal fluency	n/a
Nehra 2014	BD	euthymic	DSM-IV	n/a	outpatients	20	40.20 (10.69)	15%	20	DSST (WAIS)	n/a
Normala 2010	BD	euthymic	DSM-IV	n/a	not stated	40	Median=37.5	53%	40	TMT-A, Verbal fluency	n/a
Okasha 2014	BD	euthymic	DSM-IV	n/a	outpatients	60	27.02 (5.70)	50%	60	DSST (WAIS)	CPT (Conners)

Appendix A

Pattanayak 2012	BD	euthymic	DSM-IV	n/a	outpatients	30	33.53 (10.31)	37%	20	TMT-A, Stroop Task	n/a
Pier 2004	MDD	depressed	DSM-IV	Unmedicated.	inpatients	38	39.00 (9.00)	68%	38	DSST	n/a
Poletti 2014	BD	depressed	DSM-IV	Some reported previous psychotic symptoms.	inpatients	100	46.90 (11.91)	86%	100	DSST	n/a
Poletti 2017	BD and MDD	not stated	DSM-IV	Without psychotic features.	inpatients	133 (*n*=76 BD, *n*=57 MDD)	BD=47.75 (11.58), MDD=47.40 (10.67)	BD=67, MDD=67	57	DSST	n/a
Porter 2003	MDD	depressed	DSM-IV	Unmedicated; 26 were drug naive.	outpatients	44	32.90 (10.60)	66%	44	DSST (WAIS), Verbal fluency	CPT (Vigil)
Pradhan 2008	BD	euthymic	ICD-10 DCR	Some had psychotic symptoms.	inpatients and outpatients	48	37.23 (10.80)	19%	23	TMT-A, Verbal fluency	n/a
Radwan 2013	BD	euthymic	DSM-IV	YMRS scores: very severe >23; severe 19–22; moderate 14–18; mild 8–13; normal <7. Most had histories of predominantly manic episodes.	outpatients	30	28.67 (7.24)	50%	30	n/a	CPT
Robinson 2013	BD	euthymic	DSM-IV	n/a	outpatients	22	43.14 (7.80)	64%	21	n/a	CPT (AX)
Romero 2016	BD	euthymic	DSM-5	No history of psychosis.	outpatients	46	41.40 (18.20)	72%	46	TMT-A, Verbal fluency	n/a
Ronold 2020	MDD	depressed	DSM-IV	First episode depression of a moderate-to severe degree. Non-psychotic.	outpatients	18	27.06 (6.48)		31	TMT-A, Verbal fluency, Stroop task (D-KEFS)	n/a
Sánchez-Carro 2021	MDD	depressed	DSM-IV-TR	n/a	outpatients	74	49.25 (9.99)	75%	68	n/a	CPT (CANTAB RVIP)

Schmidt 2021	MDD	depressed	DSM-5	Non-psychotic. Unmedicated.	not stated	45	37.20 (14.17)	70%	45	n/a	CPT (CVAT)
Şentürk Cankorur 2017	BD	euthymic	DSM-IV	On monotherapy (lithium, valproate, or antipsychotics). Some with history of psychosis.	not stated	67	37.73 (12.41)		42	DSST (WAIS-R)	n/a
Shan 2011	BD	euthymic	DSM-IV-TR	Some had psychotic symptoms.	inpatients and outpatients	28	32.14 (8.96)	57%	22	DSST (WAIS-III), TMT-A	n/a
Shimizu 2013	MDD	remitted	DSM-IV	Non-psychotic.	Outpatients	43	38.30 (8.90)	23%	43	TMT-A, Verbal fluency	CPT
Smucny 2018	BD	not stated	DSM-IV-TR	Psychotic: onset of psychosis occurring within 2 years of the study.	outpatients	27	21.95 (2.72)	50%	86	n/a	CPT (AX)
Smucny 2019	BD	not stated	DSM-IV	History of psychotic features.	outpatients	58	38.66 (1.39)	67%	72	n/a	CPT (AX Dot Probe Expectancy version)
Solé 2012	BD	euthymic	DSM-IV	Some had prior psychotic symptoms.	not stated	43	46.58 (9.22)	53%	42	TMT-A, Verbal fluency	n/a
Soni 2017	BD	euthymic	ICD-10	Some had present psychotic symptoms.	outpatients	61	33.66 (9.04)	39%	30	TMT-A, Stroop task	n/a
Sun 2020	MDD	depressed	DSM-IV-TR	n/a	inpatients and outpatients	579	34.77 (12.05)	58%	321	TMT-A	n/a
Sweeney 2000	BD and MDD	depressed MDD; symptomatic (depressed, mixed, or manic) BD	DSM-IV	Some had history of psychosis.	inpatients	93	BD-m=36.14 (11.01), BD-d=31.90 (1.36), MDD=32.29 (9.10)	BD-m =57%, BD-d =43, MDD=67	51	Choice RT (Big Circle/Little Circle, Five Stage Reaction Time task)	n/a
Talarowska 2010	MDD	depressed	ICD-10	n/a	inpatients	30	45.15 (9.02)	53%	57	Stroop task	n/a

Thompson 2005	BD	euthymic	DSM-IV	n/a	outpatients	63	44.4 (8.60)	41%	63	DSST (WAIS-R), TMT-A, Verbal fluency	CPT (Vigil)
Torrent 2011	BD	euthymic	DSM-IV	n/a	outpatients	84	40.86 (12.19)	42%	35	TMT-A, Verbal fluency	n/a
Trivedi 2007	BD	euthymic	DSM-IV	n/a	outpatients	15	34.43 (10.71)	20%	15	n/a	CPT (Conners)
Vaskinn 2011	BD	not stated	DSM-IV	Some had a history of psychosis.	inpatients and outpatients	106	36.40 (11.00)	53%	340	DSST (WAIS-III), Stroop task (D-KEFS)	n/a
Vicent-Gil 2018	MDD	depressed	DSM-IV-TR	First episode.	not stated	90	43.86 (10.61)	60%	40	DSST (WAIS-III), TMT-A, Verbal fluency	CPT (II)
Walsh 2009	MDD	depressed	DSM-IV	Unmedicated.	not stated	11	39.80 (6.80)	not stated	11	DSST (WAIS)	n/a
Xu 2012	BD and MDD	depressed MDD and BD, remitted MDD and BD	DSM-IV-TR	Some had psychotic features.	inpatients and outpatients	516	BD=32.17 (11.72), MDD=34.90 (12.70)	BD=46, MDD=55	202	DSST (WAIS-R), TMT-A	n/a
Yadav 2011	BD	manic	ICD-10	n/a	not stated	50	28.54 (6.14)	0%	50	TMT-A (comprehensive TMT)	n/a
Yamamoto 2012	MDD	remitted	DSM-IV-TR	Unmedicated.	community	12	21.08 (4.62)	58%	19	TMT-A, Verbal fluency	CPT, PASAT
Zhang 2020a	BD	depressed	DSM-5	n/a	inpatients and outpatients	58	26.40 (8.59)	62%	61	THINC-it Spotter (Choice-RT), Codebreaker (DSST)	n/a
Zhang 2020b	MDD	depressed	ICD-10	Unmedicated on day of visit.	outpatients	83	38.73 (11.47)	69%	85	DSST	n/a
Zhao 2021	MDD	depressed	DSM-IV	n/a	inpatients	222	30.88 (10.95)	49%	173	TMT-A, DSST (WAIS)	n/a
Zhou 2019	MDD	depressed	DSM-5	n/a	not stated	146	34.12 (12.03)	48%	70	MCCB: TMT-A, PS composite	n/a

Zhu 2019	BD and MDD	not stated	DSM-IV	n/a	inpatients and outpatients	104	32.37 (9.26)	342%	249	MCCB: TMT-A, DSST, Verbal fluency	CPT (IP; MCCB)

Table 21: Summary of paper characteristics of included studies. Ns are those reported in the paper; note that a different N may have been used for the meta-analysis. ANT=Attention Network Test; BD=bipolar disorder; BD-e=bipolar disorder euthymic; BD-d=bipolar disorder depressed; BD-m=bipolar disorder manic; CalCAP=California Computerized Assessment Package; CANTAB=Cambridge Neuropsychological Test Automated Battery; CNSVS=CNS Vital Signs; D-KEFS=Delis-Kaplan Executive Function System; DSM=Diagnostic and Statistical Manual of Mental Disorders; DSST=Digit Symbol Substitution Test; ICD=International Classification of Diseases; MATRICS=Measurement and Treatment Research to Improve Cognition in Schizophrenia; MCCB=MATRICS Consensus Cognitive Battery; MDD=major depressive disorder; PASAT=Paced Auditory Serial Addition Test; PS=processing speed; RBANS=Repeatable Battery for the Assessment of Neuropsychological Status; RT=reaction time; SA=sustained attention; SART=Sustained Attention to Response Task; SCIP=Screen for Cognitive Impairment in Psychiatry; SCOLP=Speed and Capacity of Language Processing Test; SD=standard deviation; TMT=trail-making test; WAIS=Wechsler Adult Intelligence Scale; WASI=Wechsler Abbreviated Scale of Intelligence.

e) COUNTS OF REASONS FOR EXCLUSION

Counts	Ovid search	WoK search	Updated search	Total
Records identified from search result	17,321	17,313	1,812	36,446
After de-duplication	9,012	10,518	1,479	21,009
Included after screening titles and abstracts	1,749	394	283	2,426
Included after full-text review	78	7	18	103
Contacted authors	406	53	44	503
Included papers with MDD participants	874	168	142	1,184
Included papers with BD participants	734	176	113	1,023
Included papers with Both BD and MDD participants	141	50	28	219
Exclusion Reasons				
No healthy controls	463	89	94	646
Data missing/not available	371	51	43	465
No tests of PS or SA	275	106	37	418
Comorbidity	101	9	22	132
Redundant data	95	12	17	124
Review	85	11	4	100
Not MDD or BD (healthy controls only)	48	14	2	64
No neuropsychological measures	35	18	4	57
Diagnosis tool	51	4	0	55
Not MDD or BD (other psych disorder)	15	37	1	53
Older adults	18	5	22	45
Not original article	34	2	7	43
Imaging study	28	7	4	39
Subjective cognitive test	14	0	1	15
No MDD or BD (relatives only)	9	3	1	13
Children	7	6	0	13
Duplicate	1	11	0	12
No MDD or BD (other clinical group)	8	0	3	11
Composite data only	6	0	3	9
Eye tracking	6	2	0	8
Not English	1	0	0	1

Table 22: A count of records and exclusion reasons for each search. WoK=Web of Knowledge; BD=bipolar disorder; MDD=Major Depressive Disorder; PS=processing speed; SA=sustained attention.

f) ADDITIONAL RESULTS FROM THE META-ANALYSIS

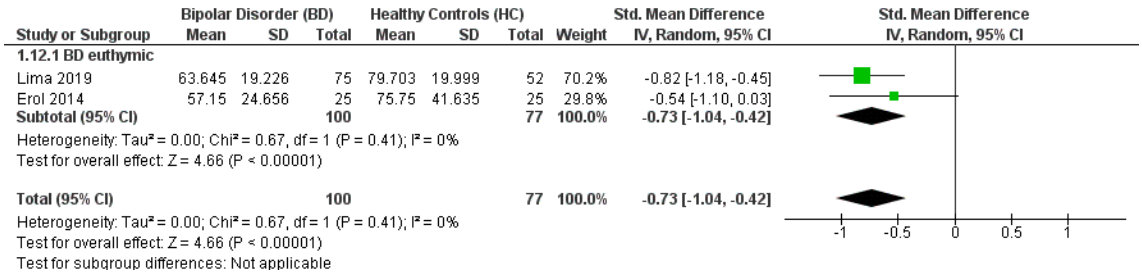

Figure 44: Forest plot of the comparison between BD patients and healthy controls on Stroop simple/automatic trials (word reading and colour naming combined) number correct.

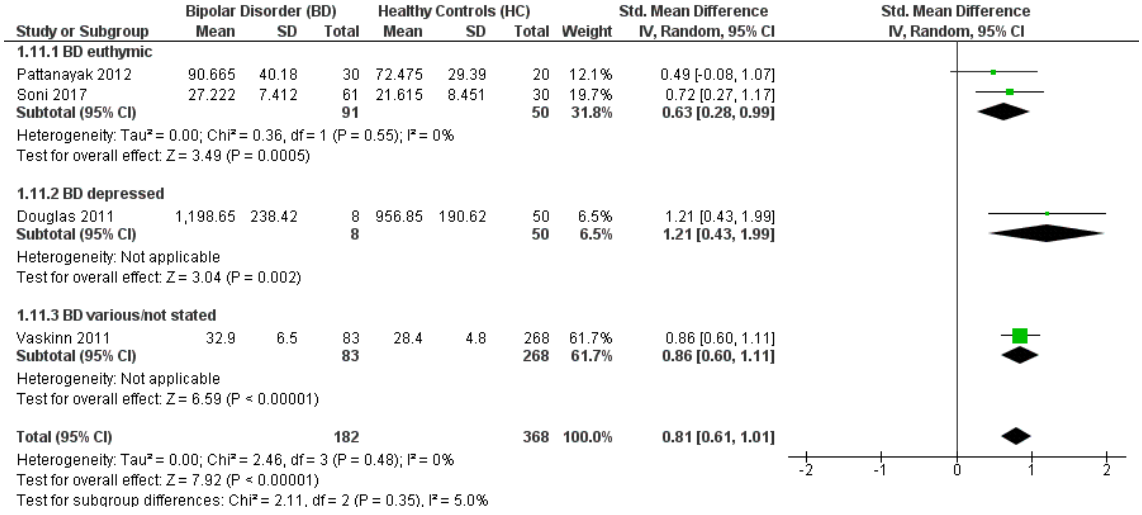

Figure 45: Forest plot of the comparison between BD patients and healthy controls on Stroop simple/automatic trials (word reading and colour naming combined) time to complete.

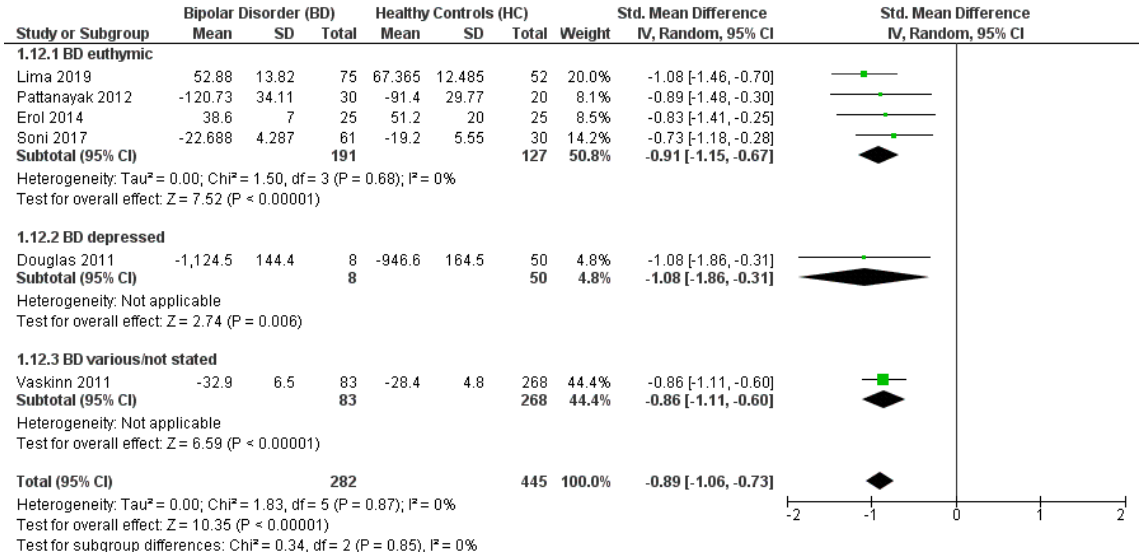

Figure 46: Forest plot of the comparison between BD patients and healthy controls on Stroop colour naming trials.

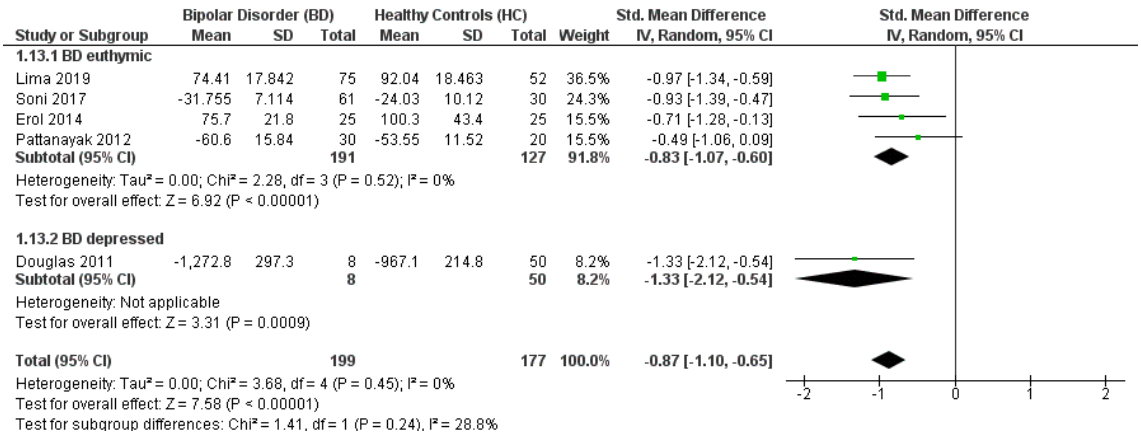

Figure 47 Forest plot of the comparison between BD patients and healthy controls on Stroop word reading trials.

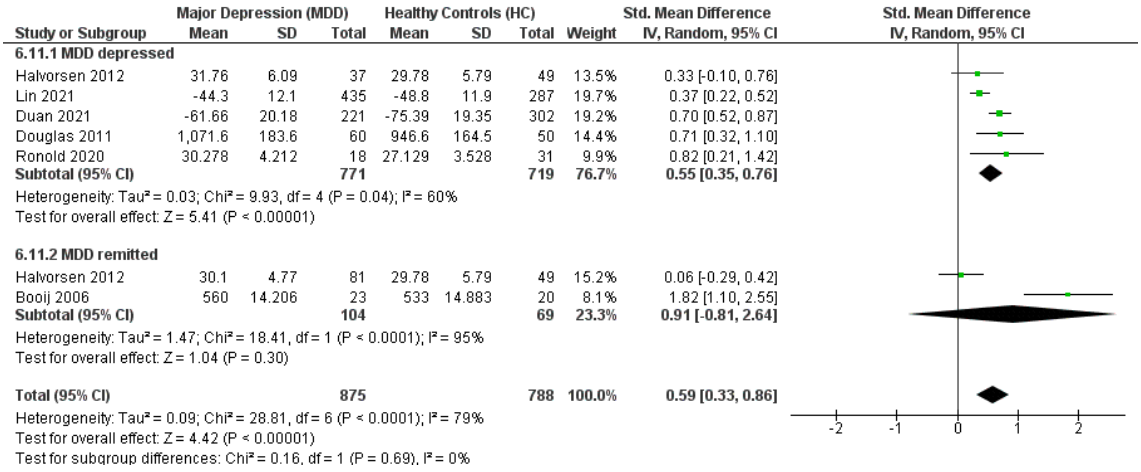

Figure 48: Forest plot of the comparison between MDD patients and healthy controls on Stroop colour naming.

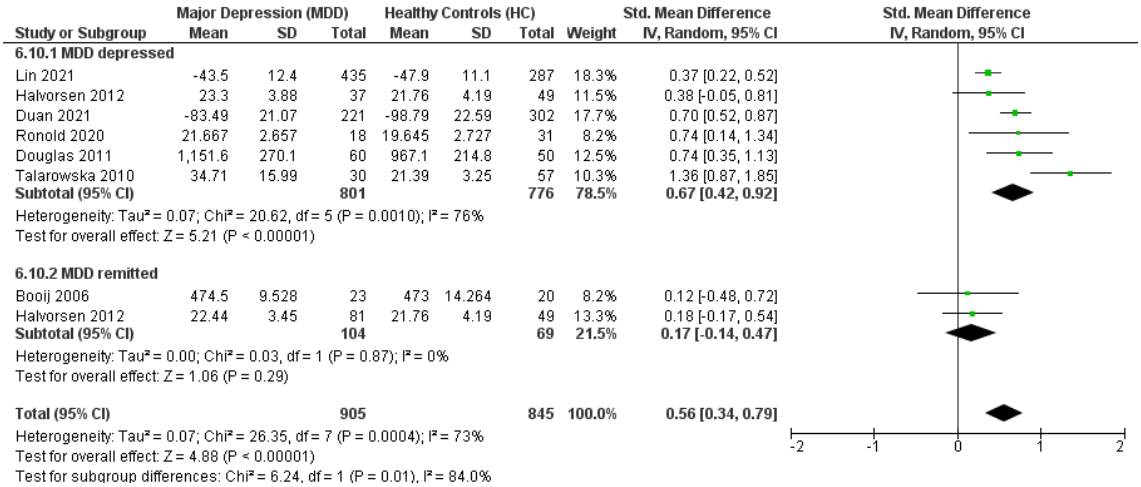

Figure 49: Forest plot of the comparison between MDD patients and healthy controls on Stroop word reading.

g) ADDITIONAL RESULTS FROM THE RISK OF BIAS ANALYSIS

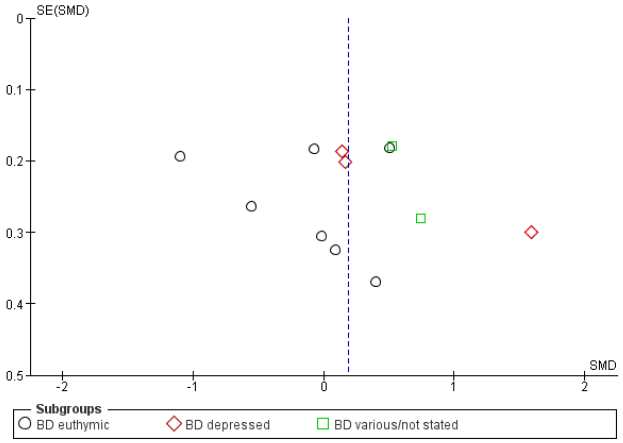

Figure 50: Funnel plot for CPT average RT for BD.

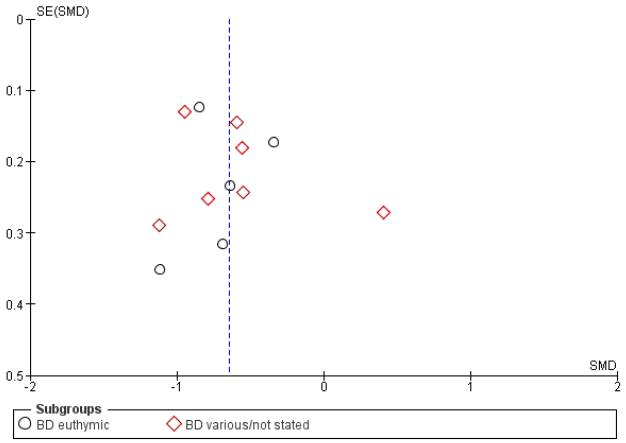

Figure 51: Funnel plot for CPT d' for BD.

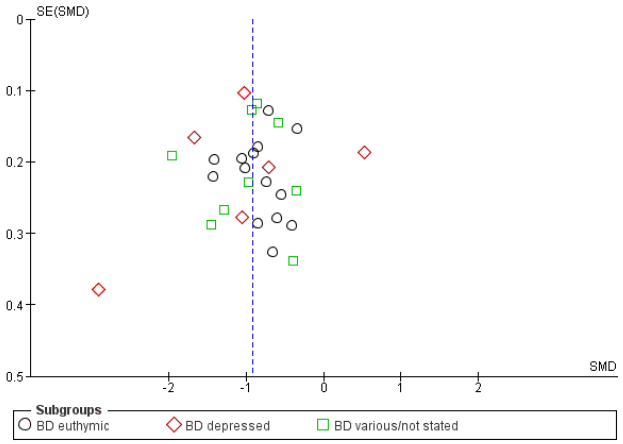

Figure 52: Funnel plot for DSST for BD.

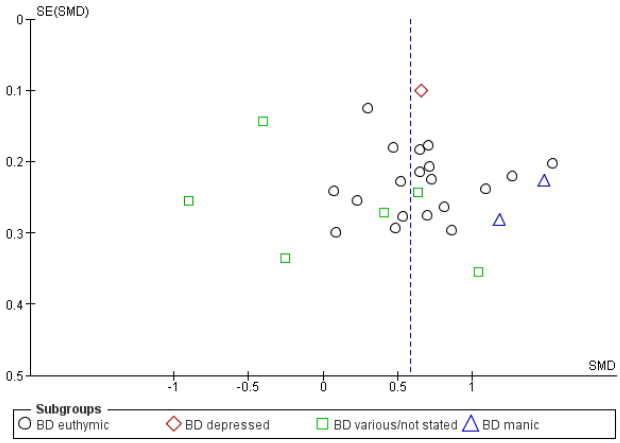

Figure 53: Funnel plot for TMT-A time for BD.

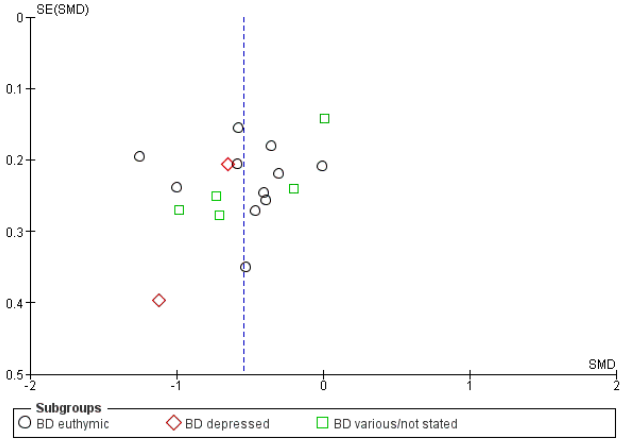

Figure 54: Funnel plot for verbal fluency for BD.

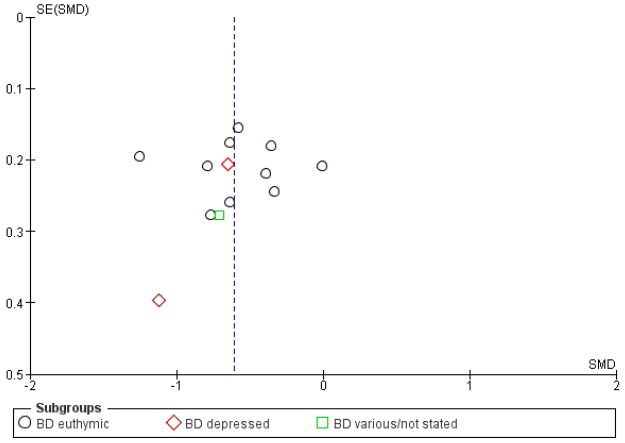

Figure 55: Funnel plot for phonemic fluency for BD.

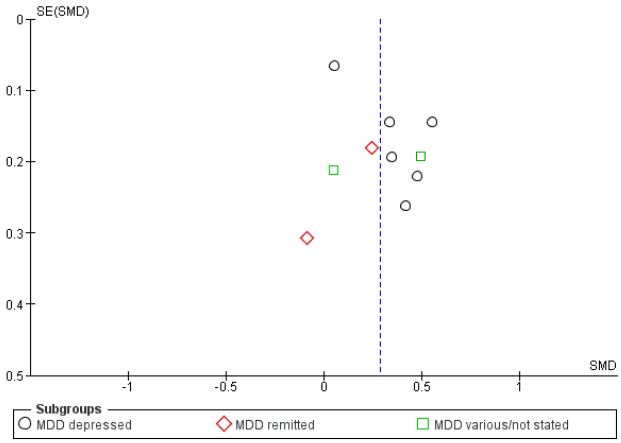

Figure 56: Funnel plot for choice-RT for MDD.

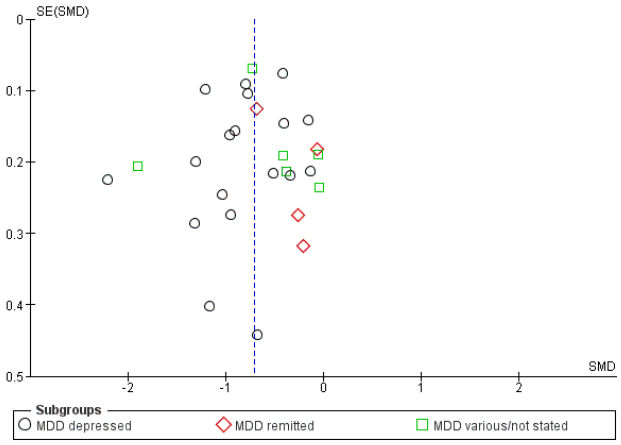

Figure 57: Funnel plot for DSST correct for MDD.

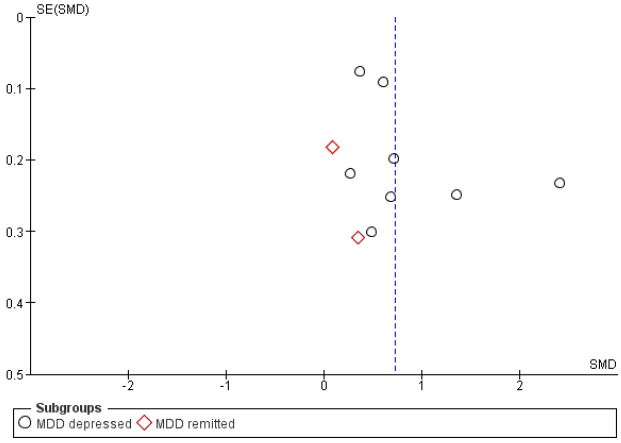

Figure 58: Funnel plot for Stroop simple trials for MDD.

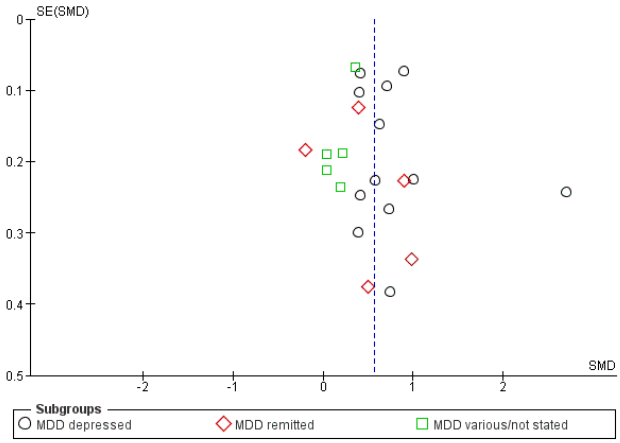

Figure 59: Funnel plot for TMT-A time MDD.

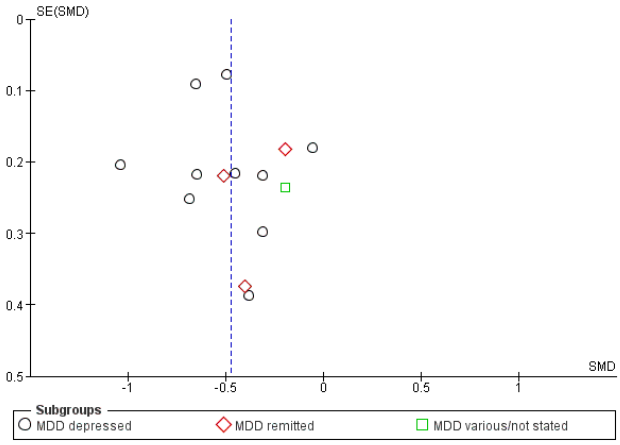

Figure 60: Funnel plot for verbal fluency for MDD.

	Bipolar disorder			Major Depressive Disorder		
	z	**p**	**intercept**	**z**	**p**	**intercept**
CPT average RT	0.84	0.40	-0.43			
CPT sensitivity (d')	0.11	0.92	-0.68			
Choice RT				0.61	0.54	0.17
DSST (correct)	-0.92	0.36	-0.61	-0.31	0.76	-0.63
Stroop simple trials				0.47	0.64	0.42
TMT-A (time to complete)	-0.03	0.97	0.59	0.76	0.45	0.38
Phonemic Fluency (correct)	-0.86	0.39	-0.27			
Verbal Fluency (correct)	-1.51	0.13	-0.02	0.55	0.58	-0.56

Table 23: Results of Egger's test of publication bias for comparisons with at least 10 studies. CPT=Continuous Performance Test; RT=reaction time; DSST=Digit Symbol Substitution Test; TMT-A=Trail-Making Test part A.

APPENDIX B

a) DETAILS OF MISSING DATA

Dataset	Percentage of missing data overall	Pattern of missing data
Bipolar disorder euthymic	1% (patients=0.6%, controls=1.8%)	NART 1%, DSST 1%, TMT-A 1%, Vigil errors 1%, Vigil *tau* 5%, Digit span 1%, TMT-B 2%, Stroop 2%, RAVLT 1%, S-Rec 0%, P-Rec 0%, Spatial Span 2%, SWM 0%
Bipolar disorder depressed	1% (patients=1%, controls=1%)	NART 0%, DSST 1%, SCOLP 2%, Vigil errors 0%, Vigil *tau* 1%, Digit span 0%, verbal fluency 0%, RAVLT 1%, S-Rec 0%, P-Rec 0%, Spatial Span 0%, SWM 0%
Major depressive disorder	3% (patients=3%, controls=3%)	NART 0%, DSST 0%, Vigil errors 0%, Vigil *tau* 16%, ToL 10%, RAVLT 0%, S-Rec 0%, P-Rec 0%, SWM 0%

Table 24: Details of missing data in each dataset. NART=National Adult Reading Test; DSST=Digit Symbol Substitution Test; TMT=Trail-Making Test; RAVLT=Rey Auditory Verbal Learning Test; S-Rec=Spatial Recognition; P-Rec=Pattern Recognition; SWM=Spatial Working Memory; ToL=Tower of London.

b) TRANSFORMATIONS OF COGNITIVE DOMAIN COMPOSITE SCORES

Dataset	Cognitive Domain	Number of outliers in control group	Transformations performed	Distribution of whole sample (after pre-processing)
Bipolar disorder euthymic	**PS**	0	^3	normally distributed
	SA	1*	^3	negatively skewed
	EF	1*	^3	normally distributed
	VM	0	^2	negatively skewed
	VS	0	reflect, square root	normally distributed
Bipolar disorder depressed	**PS**	0	square root	normally distributed
	SA	1*	^3	negatively skewed
	EF	0	^2	normally distributed
	VM	1*	^2	normally distributed
	VS	0	^3	normally distributed
Major depressive disorder	**PS**	0	n/a	normally distributed
	SA	0	^3	negatively skewed
	EF	0	n/a	normally distributed
	VM	0	^3	normally distributed
	VS	0	n/a	normally distributed

*Table 25: Details of outliers and transformations performed to the composite cognitive domain scores. Outliers and transformations are reported for the composite scores before regressing out age and premorbid IQ. Details of distributions refer to cognitive domain scores for the whole sample, after pre-processing, i.e., after outliers were Winsorized, transformations were preformed, and age and premorbid IQ were regressed out. BD=bipolar disorder; PS=processing speed; SA=sustained attention; EF=executive function; VM=verbal learning and memory; VS=visuospatial memory. *Winsorized to within 3 standard deviations of the mean.*

c) CORRELATIONS BETWEEN AGE, EDUCATION AND NART IQ

NART IQ was significantly correlated with education in BD-e patients ($r(60)=.528$, $p<.001$) and their matched control group ($r(61)=.625$, $p<.001$), BD-d patients ($r(38)=.515$, $p=.001$) and their matched control group ($r(29)=.487$, $p=.005$), and MDD patients ($r(39)=.491$, $p=.001$), but not in their matched control group ($r(38)=-.064$, $p=.693$). Age was not related to NART IQ or education in any of the samples (all $ps>.05$), except in the controls matched to MDD and BD-d, where age was correlated with NART IQ ($r(39)=.354$, $p=.023$, and $r(35)=.409$, $p=.012$, respectively).

d) RESULTS OF HIERARCHICAL REGRESSION MODELS TO TEST THE EFFECT OF CORE COGNITIVE FUNCTIONS ON THE RELATIONSHIP BETWEEN GROUP AND MEMORY FOR EACH DATASET

Verbal learning and memory for BD-e												
		Model statistics					Main effects			Change Statistics		
	Model	*F*	*df*	*p*	R^2	R^2 adj.	IV	*β*	*t*	*p*	R^2 change	*F* Change	*F*-change *p*
1	PS + Group	6.06*	2, 122	.003	.09	.08	PS	0.12	1.24	.216	-	-	-
							Group	-0.23*	-2.47	.015	-	-	-
	PS + SA + Group	5.55*	3, 121	.001	.12	.10	PS	0.07	0.73	.468	.03	-0.51	.042
							SA	0.19*	2.05	.042	-	-	-
							Group	-0.19	-1.97	.051	-	-	-
	PS + SA + EF + Group	5.97*	4, 120	<.001	.17	.14	PS	0.04	0.39	.697	.05	0.42	.012
							SA	0.11	1.11	.271	-	-	-
							EF	0.24*	2.55	.012	-	-	-
							Group	-0.16	-1.64	.104	-	-	-
2	SA + Group	8.09*	2, 122	.001	.12	.10	SA	0.22*	2.30	.023	-	-	-
							Group	-0.21*	-2.34	.021	-	-	-
	SA + PS + Group	5.55*	3, 121	.001	.12	.10	SA	0.19*	2.05	.042	<.01	-2.54	.468
							PS	0.07	0.73	.468	-	-	-
							Group	-0.19	-1.97	.051	-	-	-
	SA + PS + EF + Group	5.97*	4, 120	<.001	.17	.14	SA	0.11	1.11	.271	.05	0.42	.012
							PS	0.04	0.39	.697	-	-	-
							EF	0.24*	2.55	.012	-	-	-
							Group	-0.16	-1.64	.104	-	-	-
3	EF + Group	11.18*	2, 122	<.001	.15	.14	EF	0.29*	3.32	.001	-	-	-
							Group	-0.19*	-2.18	.031	-	-	-
	EF + PS + Group	7.54*	3, 121	<.001	.16	.14	EF	0.28*	3.11	.002	<.01	-3.64	.552
							PS	0.06	0.60	.552	-	-	-

							Group	-0.17	-1.84	.068	-	-	-
	EF + PS + SA + Group	5.97*	4, 120	<.001	.17	.14	EF	0.24*	2.55	.012	.01	-1.57	.271
							PS	0.04	0.39	.697	-	-	-
							SA	0.11	1.11	.271	-	-	-
							Group	-0.16	-1.64	.104	-	-	-

*Table 26: Detailed results of hierarchical regressions to test the effect of core cognitive functioning on the relationship between group and verbal learning and memory for the euthymic bipolar disorder dataset. *Significant at the 0.05 level. PS=processing speed; SA=sustained attention; EF=executive function; df=degrees of freedom; IV=independent variable.*

Visuo-spatial memory for BD-e													
		Model statistics					**Main effects**				**Change Statistics**		
	Model	***F***	***df***	***p***	***R²***	***R²* adj.**	**IV**	***β***	***t***	***p***	***R²* change**	***F* Change**	***F*-change *p***
1	PS + Group	7.98*	2, 122	.001	.12	.10	PS	0.29*	3.07	.003	-	-	-
							Group	0.33*	3.58	<.001	-	-	-
	PS + SA + Group	5.35*	3, 121	.002	.12	.10	PS	0.28*	2.86	.005	<.01	-2.63	.652
							SA	0.04	0.45	.652	-	-	-
							Group	0.34*	3.58	<.001	-	-	-
	PS + SA + EF + Group	6.66*	4, 120	<.001	.18	.15	PS	0.24*	2.51	.013	.06	1.30	.003
							SA	-0.06	-0.61	.544	-	-	-
							EF	0.29*	3.07	.003	-	-	-
							Group	0.38*	4.09	<.001	-	-	-
2	SA + Group	3.72*	2, 122	.027	.06	.04	SA	0.11	1.14	.255	-	-	-
							Group	0.25*	2.71	.008	-	-	-
	SA + PS + Group	5.35*	3, 121	.002	.12	.10	SA	0.04	0.45	.652	.06	1.64	.005
							PS	0.28*	2.86	.005	-	-	-
							Group	0.34*	3.58	<.001	-	-	-
	SA + PS + EF + Group	6.66*	4, 120	<.001	.18	.15	SA	-0.06	-0.61	.544	.06	1.30	.003
							PS	0.24*	2.51	.013	-	-	-
							EF	0.29*	3.07	.003	-	-	-

							Group	0.38*	4.09	<.001	-	-	-
3	EF + Group	9.79*	2, 122	<.001	.14	.12	EF	0.32*	3.59	<.001	-	-	-
							Group	0.31*	3.57	.001	-	-	-
	EF + PS + Group	8.80*	3, 121	<.001	.18	.16	EF	0.27*	3.06	.003	.04	-0.99	.016
							PS	0.23*	2.45	.016	-	-	-
							Group	0.39*	4.26	<.001	-	-	-
	EF + PS + SA + Group	6.66*	4, 120	<.001	.18	.15	EF	0.29*	3.07	.003	<.01	-2.14	.544
							PS	0.24*	2.51	.013	-	-	-
							SA	-0.06	-0.61	.544	-	-	-
							Group	0.38*	4.09	<.001	-	-	-

*Table 27: Detailed results of hierarchical regressions to test the effect of core cognitive functioning on the relationship between group and visuo-spatial memory for the euthymic bipolar disorder dataset. *Significant at the 0.05 level. PS=processing speed; SA=sustained attention; EF=executive function; df=degrees of freedom; IV=independent variable.*

Verbal learning and memory for BD-d													
		Model statistics					**Main effects**				**Change Statistics**		
	Model	***F***	***df***	***p***	***R^2***	***R^2* adj.**	**IV**	***β***	***t***	***p***	***R^2* change**	***F* Change**	***F*-change *p***
1	PS + Group	1.84	2, 78	.166	.05	.02	PS	14.16	1.03	.305	-	-	-
							Group	4.08	-0.80	.429	-	-	-
	PS + SA + Group	1.59	3, 77	.199	.06	.02	PS	14.22	0.93	.356	.01	-0.25	.302
							SA	0.01	1.04	.302	-	-	-
							Group	4.27	-0.45	.653	-	-	-
	PS + SA + EF + Group	1.39	4, 76	.247	.07	.02	PS	14.23	0.93	.355	.01	-0.20	.375
							SA	0.01	0.85	.400	-	-	-
							EF	0.13	0.89	.375	-	-	-
							Group	4.28	-0.50	.616	-	-	-
2	SA + Group	1.95	2, 78	.149	.05	.02	SA	0.01	1.13	.261	-	-	-
							Group	3.72	-1.04	.302	-	-	-
	SA + PS + Group	1.59	3, 77	.199	.06	.02	SA	0.01	1.04	.302	.01	-0.37	.356

							PS	14.22	0.93	.356	-	-	-
							Group	4.27	-0.45	.653	-	-	-
	SA + PS + EF + Group	1.39	4, 76	.247	.07	.02	SA	0.01	0.85	.400	.01	-0.20	.375
							PS	14.23	0.93	.355	-	-	-
							EF	0.13	0.89	.375	-	-	-
							Group	4.28	-0.50	.616	-	-	-
3	EF + Group	1.91	2, 78	.156	.05	.02	EF	0.13	1.09	.279	-	-	-
							Group	3.42	-1.61	.111	-	-	-
	EF + PS + Group	1.62	3, 77	.192	.06	.02	EF	0.13	1.08	.285	.01	-0.29	.312
							PS	14.14	1.02	.312	-	-	-
							Group	4.07	-0.80	.427	-	-	-
	EF + PS + SA + Group	1.39	4, 76	.247	.07	.02	EF	0.13	0.89	.375	.01	-0.23	.400
							PS	14.23	0.93	.355	-	-	-
							SA	0.01	0.85	.400	-	-	-
							Group	4.28	-0.50	.616	-	-	-

*Table 28: Detailed results of hierarchical regressions to test the effect of core cognitive functioning on the relationship between group and verbal learning memory for the depressed bipolar disorder dataset. *Significant at the 0.05 level. PS=processing speed; SA=sustained attention; EF=executive function; df=degrees of freedom; IV=independent variable.*

Visuo-spatial memory for BD-d												
		Model statistics					**Main effects**			**Change Statistics**		
	Model	***F***	***df***	***p***	***R²***	***R²* adj.**	**IV**	***β***	***t***	***p***	***R²* change**	***F* Change**	***F*-change *p***
1	PS + Group	2.12	2, 78	.127	.05	.03	PS	169.15	1.70	.094	-	-	-
							Group	48.71	-0.06	.955	-	-	-
	PS + SA + Group	6.00*	3, 77	.001	.19	.16	PS	158.10	1.47	.146	.14	3.88	.001
							SA	0.07*	3.62	.001	-	-	-
							Group	47.46	1.02	.312	-	-	-
	PS + SA + EF + Group	5.27*	4, 76	.001	.22	.18	PS	156.39	1.49	.140	.03	-0.73	.105
							SA	0.07*	3.28	.002	-	-	-

							EF	1.44	1.64	.105	-	-	-
							Group	47.03	0.93	.356	-	-	-
2	SA + Group	7.80*	2, 78	.001	.17	.15	SA	0.07*	3.75	<.001	-	-	-
							Group	41.71	0.34	.735	-	-	-
	SA + PS + Group	6.00*	3, 77	.001	.19	.16	SA	0.07*	3.62	.001	.02	-1.81	.146
							PS	158.10	1.47	.146	-	-	-
							Group	47.46	1.02	.312	-	-	-
	SA + PS + EF + Group	5.27*	4, 76	.001	.22	.18	SA	0.07*	3.28	.002	.03	-0.73	.105
							PS	156.39	1.49	.140	-	-	-
							EF	1.44	1.64	.105	-	-	-
							Group	47.03	0.93	.356	-	-	-
3	EF + Group	3.07	2, 78	.052	.07	.05	EF	1.52*	2.18	.032	-	-	-
							Group	40.40	-1.16	.248	-	-	-
	EF + PS + Group	3.07*	3, 77	.033	.11	.07	EF	1.50*	2.18	.032	.03	-0.01	.092
							PS	165.24	1.70	.093	-	-	-
							Group	47.58	-0.06	.951	-	-	-
	EF + PS + SA + Group	5.27*	4, 76	.001	.22	.18	EF	1.44	1.64	.105	.11	2.21	.002
							PS	156.39	1.49	.140	-	-	-
							SA	0.07*	3.28	.002	-	-	-
							Group	47.03	0.93	.356	-	-	-

*Table 29: Detailed results of hierarchical regressions to test the effect of core cognitive functioning on the relationship between group and visuo-spatial memory for the depressed bipolar disorder dataset. *Significant at the 0.05 level. PS=processing speed; SA=sustained attention; EF=executive function; df=degrees of freedom; IV=independent variable.*

Verbal learning and memory for MDD													
		Model statistics					**Main effects**				**Change Statistics**		
	Model	***F***	***df***	***p***	***R^2***	***R^2* adj.**	**IV**	***β***	***t***	***p***	***R^2* change**	***F* Change**	***F*-change *p***
1	PS + Group	2.56	2, 79	.084	.06	.04	PS	27.10	-0.10	.923	-	-	-
							Group	47.57*	-2.26	.027	-	-	-

	PS + SA + Group	2.31	3, 78	.083	.08	.05	PS	28.17	-0.48	.636	.02	-0.25	.189
							SA	0.08	1.33	.189	-	-	-
							Group	53.05	-1.43	.157	-	-	-
	PS + SA + EF + Group	2.06	4, 77	.094	.10	.05	PS	28.22	-0.57	.568	.02	-0.25	.259
							SA	0.08	1.42	.161	-	-	-
							EF	25.37	-1.14	.259	-	-	-
							Group	53.27	-1.30	.198	-	-	-
2	SA + Group	3.38*	2, 79	.039	.08	.06	SA	0.08	1.25	.216	-	-	-
							Group	52.67	-1.47	.145	-	-	-
	SA + PS + Group	2.31	3, 78	.083	.08	.05	SA	0.08	1.33	.189	<.01	-1.08	.636
							PS	28.17	-0.48	.636	-	-	-
							Group	53.05	-1.43	.157	-	-	-
	SA + PS + EF + Group	2.06	4, 77	.094	.10	.05	SA	0.08	1.42	.161	.02	-0.25	.259
							PS	28.22	-0.57	.568	-	-	-
							EF	25.37	-1.14	.259	-	-	-
							Group	53.27	-1.30	.198	-	-	-
3	EF + Group	3.11	2, 79	.050	.07	.05	EF	25.23	-1.02	.312	-	-	-
							Group	47.31*	-2.18	.032	-	-	-
	EF + PS + Group	2.05	3, 78	.113	.07	.04	EF	25.45	-1.02	.311	<.01	-1.05	.870
							PS	27.16	-0.16	.870	-	-	-
							Group	47.71*	-2.17	.033	-	-	-
	EF + PS + SA + Group	2.06	4, 77	.094	.10	.05	EF	25.37	-1.14	.259	.02	0.01	.161
							PS	28.22	-0.57	.568	-	-	-
							SA	0.08	1.42	.161	-	-	-
							Group	53.27	-1.30	.198	-	-	-

*Table 30: Detailed results of hierarchical regressions to test the effect of core cognitive functioning on the relationship between group and verbal learning and memory for the major depressive disorder dataset. *Significant at the 0.05 level. PS=processing speed; SA=sustained attention; EF=executive function; df=degrees of freedom; IV=independent variable.*

Appendix B

Visuo-spatial memory for MDD													
		Model statistics					Main effects				Change Statistics		
	Model	*F*	*df*	*p*	R^2	R^2 adj.	IV	*β*	*t*	*p*	R^2 change	*F* Change	*F*-change *p*
1	PS + Group	8.25*	2, 79	.001	.17	.15	PS	0.098	1.64	.105	-	-	-
							Group	0.173*	-3.59	.001	-	-	-
	PS + SA + Group	7.88*	3, 78	<.001	.23	.20	PS	0.100	0.91	.365	.06	-0.37	.016
							SA	<.001*	2.46	.016	-	-	-
							Group	0.188*	-2.19	.032	-	-	-
	PS + SA + EF + Group	5.92*	4, 77	<.001	.24	.20	PS	0.100	0.95	.346	<.01	-1.96	.612
							SA	<.001*	2.40	.019	-	-	-
							EF	0.090	0.51	.612	-	-	-
							Group	0.190*	-2.22	.029	-	-	-
2	SA + Group	11.42*	2, 79	<.001	.22	.20	SA	<.001*	2.85	.006	-	-	-
							Group	0.187*	-2.13	.036	-	-	-
	SA + PS + Group	7.88*	3, 78	<.001	.23	.20	SA	<.001*	2.46	.016	.01	-3.54	.365
							PS	0.100	0.91	.365	-	-	-
							Group	0.188*	-2.19	.032	-	-	-
	SA + PS + EF + Group	5.92*	4, 77	<.001	.24	.20	SA	<.001*	2.40	.019	<.01	-1.96	.612
							PS	0.100	0.95	.346	-	-	-
							EF	0.090	0.51	.612	-	-	-
							Group	0.190*	-2.22	.029	-	-	-
3	EF + Group	6.86*	2, 79	.002	.15	.13	EF	0.094	0.57	.572	-	-	-
							Group	0.175*	-3.70	<.001	-	-	-
	EF + PS + Group	5.62*	3, 78	.002	.18	.15	EF	0.093	0.68	.496	.03	-1.24	.097
							PS	0.099	1.68	.097	-	-	-
							Group	0.174*	-3.62	.001	-	-	-
	EF + PS + SA + Group	5.92*	4, 77	<.001	.24	.20	EF	0.090	0.51	.612	.06	0.30	.019
							PS	0.100	0.95	.346	-	-	-
							SA	<.001*	2.40	.019	-	-	-

Group	0.190*	-2.22	.029	-	-	-

*Table 31: Detailed results of hierarchical regressions to test the effect of core cognitive functioning on the relationship between group and visuo-spatial memory for the major depressive disorder dataset. *Significant at the 0.05 level. PS=processing speed; SA=sustained attention; EF=executive function; df=degrees of freedom; IV=independent variable.*

e) RESULTS OF MULTIPLE REGRESSION MODELS TO TEST THE EFFECT OF CORE COGNITIVE FUNCTIONS ON THE RELATIONSHIP BETWEEN GROUP AND MEMORY FOR EACH DATASET

		Verbal learning and memory								Visuo-spatial memory							
BD euthymic	**Patients**	***F***	***df***	***p***	***R^2***	***R^2 adj.***	**IV**	***β***	***t***	***p***	***F***	***df***	***p***	***R^2***	***R^2 adj.***	**IV**	***β***	***t***	***p***
		3.73*	3, 58	.016	.16	.12	PS	0.08	0.66	.515	4.80*	3, 58	.005	.20	.16	PS	0.29*	2.37	.021
							SA	0.10	0.71	.478						SA	-0.06	-0.48	.636
							EF	0.31*	2.30	.025						EF	0.29*	2.19	.033
	Controls	**Verbal learning and memory**									**Visuo-spatial memory**								
		F	***df***	***p***	***R^2***	***R^2 adj.***	**IV**	***β***	***t***	***p***	***F***	***df***	***p***	***R^2***	***R^2 adj.***	**IV**	***β***	***t***	***p***
		0.73	3, 59	.540	.04	-.01	PS	-0.04	-0.34	.737	1.60	3, 59	.200	.08	.03	PS	0.09	0.72	.473
							SA	0.08	0.63	.534						SA	-0.07	-0.54	.592
							EF	0.15	1.14	.260						EF	0.26*	2.01	.049
BD depressed	**Patients**	**Verbal learning and memory**									**Visuo-spatial memory**								
		F	***df***	***p***	***R^2***	***R^2 adj.***	**IV**	***β***	***t***	***p***	***F***	***df***	***p***	***R^2***	***R^2 adj.***	**IV**	***β***	***t***	***p***
		0.47	3, 39	.702	.04	-.04	PS	-0.06	-0.37	.717	2.17	3, 39	.107	.14	.08	PS	0.16	1.07	.293
							SA	0.18	1.15	.257						SA	0.25	1.66	.105
							EF	0.02	0.12	.904						EF	0.17	1.14	.260
	Controls	**Verbal learning and memory**									**Visuo-spatial memory**								
		F	***df***	***p***	***R^2***	***R^2 adj.***	**IV**	***β***	***t***	***p***	***F***	***df***	***p***	***R^2***	***R^2 adj.***	**IV**	***β***	***t***	***p***
		2.36*	3, 34	.089	.17	.10	PS	0.34*	2.18	.036	6.74*	3, 34	.001	.37	.32	PS	0.17	1.26	.217
							SA	-0.06	-0.35	.729						SA	0.55*	3.80	<.001
							EF	0.27	1.639	.110						EF	0.10	0.66	.513
MDD	**Patients**	**Verbal learning and memory**									**Visuo-spatial memory**								
		F	***df***	***p***	***R^2***	***R^2 adj.***	**IV**	***β***	***t***	***p***	***F***	***df***	***p***	***R^2***	***R^2 adj.***	**IV**	***β***	***t***	***p***
		1.41	3, 37	.255	.10	.03	PS	-0.02	-0.16	.878	1.58	3, 37	.210	.11	.04	PS	0.20	1.27	.212
							SA	0.26	1.67	.104						SA	0.25	1.60	.118
							EF	-0.20	-1.26	.217						EF	0.01	0.09	.926

	Verbal learning and memory									Visuo-spatial memory								
	F	***df***	***p***	***R^2***	***R^2 adj.***	**IV**	***β***	***t***	***p***	***F***	***df***	***p***	***R^2***	***R^2 adj.***	**IV**	***β***	***t***	***p***
Controls	0.07	3, 37	.978	.01	-.08	PS	-0.06	-0.31	.761	2.66	3, 37	.062	.18	.11	PS	-0.14	-0.75	.458
						SA	0.21	0.09	.929						SA	0.46*	2.45	.019
						EF	-0.05	-0.30	.768						EF	0.12	0.77	.444

Table 32: Results of multiple regression models to test whether core cognitive functions can explain memory performance in patients and control groups for each dataset. Models were run separately for verbal learning and memory and visuo-spatial memory as the dependent variable. df=degrees of freedom; PS=processing speed; SA=sustained attention; EF=executive function. BD=bipolar disorder; MDD=major depressive disorder.

APPENDIX C

a) DESCRIPTIVE STATISTICS AND GROUP DIFFERENCES ON NEUROPSYCHOLOGICAL PERFORMANCE AFTER IMPUTING MISSING DATA AND CONTROLLING FOR AGE AND PREMORBID IQ

	Healthy Controls				Bipolar Disorder				Group differences			
	N	Mean	SD	Median	N	Mean	SD	Median	Statistic (*t*/*W*)	df	p-value	Cohen's *d*
d2 Concentration Performance	26	0	1	-0.18	56	-0.72	1.31	-0.83	-2.73*	62.47	.008	0.595
d2 Percent Error	26	0	1	0.19	56	-1.29	2.44	-1.11	504*	-	.026	0.625
d2 Fluctuation Rate	26	0	1	-0.06	56	-0.44	1.88	-0.24	626	-	.312	0.268
DSST Symbol minus Copy time	26	0	1	-0.02	56	-1.81	2.81	-1.30	372*	-	<.001	0.766
TMT-A time	26	0	1	0.22	56	-0.49	1.36	-0.11	590	-	.171	0.397
TMT-B minus TMT-A time	26	0	1	-0.02	56	-0.71	1.57	-0.31	541	-	.063	0.509
Category Fluency total	26	0	1	0.03	56	-0.6	1.39	-0.64	-2.22*	66.07	.030	0.474

*Table 33: Descriptive statistics and group differences on neuropsychological variables after missing data were imputed and age and premorbid IQ (NART score) were regressed out. *Significant at the 0.05 level. SD=standard deviation; df=degrees of freedom; DSST=Digit Symbol Substitution Test; TMT=Trail-Making Test.*

REFERENCES

Abé, C., Rolstad, S., Petrovic, P., Ekman, C. J., Sparding, T., Ingvar, M., & Landén, M. (2018). Bipolar disorder type I and II show distinct relationships between cortical thickness and executive function. *Acta Psychiatrica Scandinavica*, *138*(4), 325–335. https://doi.org/10.1111/acps.12922

Ahern, E., & Semkovska, M. (2017). Cognitive functioning in the first-episode of major depressive disorder: A systematic review and meta-analysis. *Neuropsychology*, *31*(1), 52–72. https://doi.org/10.1037/neu0000319

Ajilore, O., Vizueta, N., Walshaw, P., Zhan, L., Leow, A., & Altshuler, L. L. (2015). Connectome signatures of neurocognitive abnormalities in euthymic bipolar I disorder. *Journal of Psychiatric Research*, *68*, 37–44. https://doi.org/10.1016/j.jpsychires.2015.05.017

Alonso-Lana, S., Goikolea, J. M., Bonnin, C. M., Sarró, S., Segura, B., Amann, B. L., Monté, G. C., Moro, N., Fernandez-Corcuera, P., Maristany, T., Salvador, R., Vieta, E., Pomarol-Clotet, E., & McKenna, P. J. (2016). Structural and functional brain correlates of cognitive impairment in euthymic patients with bipolar disorder. *PLoS ONE*, *11*(7), 1–17. https://doi.org/10.1371/journal.pone.0158867

Alves, P. N., Foulon, C., Karolis, V., Bzdok, D., Margulies, D. S., Volle, E., & Thiebaut de Schotten, M. (2019). An improved neuroanatomical model of the default-mode network reconciles previous neuroimaging and neuropathological findings. *Communications Biology*, *2*(1), 1–14. https://doi.org/10.1038/s42003-019-0611-3

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders (4th edn) (DSM-IV)* (4th ed.). APA.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders, Fifth Edition (DSM-5)* (5th ed.). American Psychiatric Publishing, Inc. https://doi.org/10.1176/appi.books.9780890425596

Anaya, C., Martinez Aran, A., Ayuso-Mateos, J. L., Wykes, T., Vieta, E., & Scott, J. (2012). A systematic review of cognitive remediation for schizo-affective and affective disorders. *Journal of Affective Disorders*, *142*(1–3), 13–21. https://doi.org/10.1016/j.jad.2012.04.020

Ancín, I., Santos, J. L., Teijeira, C., Sánchez-Morla, E. M., Bescós, M. J., Argudo, I., Torrijos, S., Vázquez-Álvarez, B., De La Vega, I., López-Ibor, J. J., Barabash, A., & Cabranes-Díaz, J. A. (2010). Sustained attention as a potential endophenotype for bipolar disorder. *Acta Psychiatrica Scandinavica*, *122*(3), 235–245. https://doi.org/10.1111/j.1600-0447.2009.01532.x

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*(2), 71–82. https://doi.org/10.1076/chin.8.2.71.8724

Ang, Y. S., Frontero, N., Belleau, E., & Pizzagalli, D. A. (2020). Disentangling vulnerability, state and trait features of neurocognitive impairments in depression. *Brain*, *143*(12), 3865–3877. https://doi.org/10.1093/brain/awaa314

Angst, J., Gamma, A., Bowden, C. L., Azorin, J. M., Perugi, G., Vieta, E., & Young, A. H. (2013). Evidence-based definitions of bipolar-I and bipolar-II disorders among 5,635 patients with major depressive episodes in the Bridge Study: Validity and comorbidity. *European Archives of Psychiatry and Clinical Neuroscience*, *263*(8), 663–673. https://doi.org/10.1007/s00406-013-0393-4

Antila, M., Kieseppä, T., Partonen, T., Lönnqvist, J., & Tuulio-Henriksson, A. (2011). The effect of processing speed on cognitive functioning in patients with familial bipolar I disorder and their unaffected relatives. *Psychopathology*, *44*(1), 40–45. https://doi.org/10.1159/000317577

Arsintescu, L., Kato, K. H., Cravalho, P. F., Feick, N. H., Stone, L. S., & Flynn-Evans, E. E. (2019). Validation of a touchscreen psychomotor vigilance task. *Accident Analysis and Prevention*, *126*, 173–176. https://doi.org/10.1016/j.aap.2017.11.041

Arslan, F. C., Tiryaki, A., & Ozkorumak, E. (2014). A comparison of euthymic bipolar patients with unaffected first-degree relatives and healthy controls in terms of neuropsychological functions. *International Journal of Psychiatry in Clinical Practice*, *18*(3), 208–214. https://doi.org/10.3109/13651501.2013.859706

Arts, B., Jabben, N., Krabbendam, L., & Van Os, J. (2008). Meta-analyses of cognitive functioning in euthymic bipolar patients and their first-degree relatives. *Psychological Medicine*, *38*(6), 771–785. https://doi.org/10.1017/S0033291707001675

Austin, M. P., Mitchell, P., & Goodwin, G. M. (2001). Cognitive deficits in depression: Possible implications for functional neuropathology. *British Journal of Psychiatry*, *178*, 200–206. https://doi.org/10.1192/bjp.178.3.200

Azur, M. J., Stuart, E. A., Frangakis, C., & Leaf, P. J. (2011). Multiple imputation by chained equations: What is it and how does it work? *International Journal of Methods in Psychiatric Research*, *20*(1), 40–49. https://doi.org/10.1002/mpr.329

Baddeley, A. D., Emslie, H., & Nimmo-Smith, I. (1992). *The Speed and Capacity of Language Processing (SCOLP) Test*. Thames Valley Test Company.

Bakusic, J., Vrieze, E., Ghosh, M., Pizzagalli, D. A., Bekaert, B., Claes, S., & Godderis, L. (2021). Interplay of Val66Met and BDNF methylation: effect on reward learning and cognitive performance in major depression. *Clinical Epigenetics*, *13*(14), 1–11. https://doi.org/10.1186/s13148-021-01136-z

Barcikowski, R. S., & Stevens, J. P. (1975). A monte carlo study of the stability of canonical correlations, canonical weights and canonical variate-variable correlations. *Multivariate Behavioral Research*, *10*(3), 353–364. https://doi.org/10.1207/s15327906mbr1003_8

Barlow, D. H., Allen, L. B., & Choate, M. L. (2016). Toward a Unified Treatment for Emotional Disorders – Republished Article. *Behavior Therapy*, *47*(6), 838–853. https://doi.org/10.1016/j.beth.2016.11.005

Baron, R. M. R. M., & Kenny, D. A. D. A. (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Bates, M. E., & Lemay, E. P. (2004). The d2 test of attention: Construct validity and extensions in scoring techniques. *Journal of the International Neuropsychological Society*, *10*(3), 392–400. https://doi.org/10.1017/S135561770410307X

Baune, B. T., Brignone, M., & Larsen, K. G. (2018). A network meta-analysis comparing effects of various antidepressant classes on the digit symbol substitution test (DSST) as a measure of cognitive dysfunction in patients with major depressive disorder. *International Journal of Neuropsychopharmacology*, *21*(2), 97–107. https://doi.org/10.1093/ijnp/pyx070

Baune, B. T., & Malhi, G. S. (2015). A review on the impact of cognitive dysfunction on social, occupational, and general functional outcomes in bipolar disorder. *Bipolar Disorders*, *17*(Suppl. 2), 41–55. https://doi.org/10.1111/bdi.12341

Behnken, A., Bellingrath, S., Symanczik, J. P., Rieck, M. J., Zavorotnyy, M., Domschke, K., Arolt, V., & Zwanzger, P. (2013). Associations between cognitive performance and cortisol reaction to the DEX/CRH test in patients recovered from depression. *Psychoneuroendocrinology*, *38*(3), 447–454. https://doi.org/10.1016/j.psyneuen.2012.07.005

Bhardwaj, A., Wilkinson, P., Srivastava, C., & Sharma, M. (2010). Cognitive deficits in euthymic patients with recurrent depression. *Journal of Nervous and Mental Disease*, *198*(7), 513–515. https://doi.org/10.1097/NMD.0b013e3181e4c5ba

Bo, Q., Mao, Z., Li, X., Wang, Z., Wang, C., & Ma, X. (2017). Use of the MATRICS consensus cognitive battery (MCCB) to evaluate cognitive deficits in bipolar disorder: A systematic review and meta-analysis. *PLoS ONE*, *12*(4), e0176212. https://doi.org/10.1371/journal.pone.0176212

Bonnín, C. M., Martínez-Arán, A., Torrent, C., Pacchiarotti, I., Rosa, A. R., Franco, C., Murru, A., Sanchez-Moreno, J., & Vieta, E. (2010). Clinical and neurocognitive predictors of functional outcome in bipolar euthymic patients: A long-term, follow-up study. *Journal of Affective Disorders*, *121*(1–2), 156–160. https://doi.org/10.1016/j.jad.2009.05.014

Booij, L., Merens, W., Markus, C. R., & Van Der Does, A. J. W. (2006). Diet rich in α-lactalbumin improves memory in unmedicated recovered depressed patients and matched controls. *Journal of Psychopharmacology*, *20*(4), 526–535. https://doi.org/10.1177/0269881105057466

Bora, E. (2018). Neurocognitive features in clinical subgroups of bipolar disorder: A meta-analysis. *Journal of Affective Disorders*, *229*(1606), 125–134. https://doi.org/10.1016/j.jad.2017.12.057

Bora, E., Harrison, B. J., Yücel, M., & Pantelis, C. (2013). Cognitive impairment in euthymic major depressive disorder: A meta-analysis. *Psychological Medicine*, *43*(10), 2017–2026. https://doi.org/10.1017/S0033291712002085

Bora, E., Vahip, S., & Akdeniz, F. (2006). Sustained attention deficits in manic and euthymic patients with bipolar disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *30*(6), 1097–1102. https://doi.org/10.1016/j.pnpbp.2006.04.016

Bora, E., Vahip, S., Akdeniz, F., Gonul, A. S., Eryavuz, A., Ogut, M., & Alkan, M. (2007). The effect of previous psychotic mood episodes on cognitive impairment in euthymic bipolar patients. *Bipolar Disorders*, *9*(5), 468–477. https://doi.org/10.1111/j.1399-5618.2007.00469.x

Bora, E., Yucel, M., & Pantelis, C. (2009). Cognitive endophenotypes of bipolar disorder: A meta-analysis of neuropsychological deficits in euthymic patients and their first-degree relatives. *Journal of Affective Disorders*, *113*, 1–20. https://doi.org/10.1016/j.jad.2008.06.009

Bora, E., Yücel, M., & Pantelis, C. (2010). Neurocognitive markers of psychosis in bipolar disorder: A meta-analytic study. *Journal of Affective Disorders*, *127*(1–3), 1–9. https://doi.org/10.1016/j.jad.2010.02.117

Bortolato, B., Miskowiak, K. W., Köhler, C. A., Vieta, E., & Carvalho, A. F. (2015). Cognitive dysfunction in bipolar disorder and schizophrenia: A systematic review of meta-analyses. *Neuropsychiatric Disease and Treatment*, *11*, 3111–3125. https://doi.org/10.2147/NDT.S76700

Bourne, C., Aydemir, O., Balanzá-Martínez, V., Bora, E., Brissos, S., Cavanagh, J. T. O., Clark, L., Cubukcuoglu, Z., Dias, V. V., Dittmann, S., Ferrier, I. N., Fleck, D. E., Frangou, S., Gallagher, P., Jones, L., Kieseppä, T., Martínez-Arán, A., Melle, I., Moore, P. B., … Martínez-Aran, A. (2013). Neuropsychological testing of cognitive impairment in euthymic bipolar disorder: An individual patient data meta-analysis. *Acta Psychiatrica Scandinavica*, *128*(3), 149–162. https://doi.org/10.1111/acps.12133

Box, G. E. P., & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, *26*(2), 211–252. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Bradley, A. J., Anderson, K. N., Gallagher, P., & McAllister-Williams, R. H. (2019). The association between sleep and cognitive abnormalities in bipolar disorder. *Psychological Medicine*, 1–8. https://doi.org/10.1017/S0033291718004038

Bramer, W. M., Rethlefsen, M. L., Kleijnen, J., & Franco, O. H. (2017). Optimal database combinations

for literature searches in systematic reviews: A prospective exploratory study. *Systematic Reviews*, *6*(1), 1–12. https://doi.org/10.1186/s13643-017-0644-y

Braund, T. A., Tillman, G., Palmer, D. M., & Harris, A. W. F. (2020). Verbal memory predicts treatment outcome in syndromal anxious depression: An iSPOT-D report. *Journal of Affective Disorders*, *260*(September 2019), 245–253. https://doi.org/10.1016/j.jad.2019.09.028

Bromet, E., Andrade, L. H., Hwang, I., Sampson, N. A., Alonso, J., de Girolamo, G., de Graaf, R., Demyttenaere, K., Hu, C., Iwata, N., Karam, A. N., Kaur, J., Kostyuchenko, S., Lépine, J. P., Levinson, D., Matschinger, H., Mora, M. E. M., Browne, M. O., Posada-Villa, J., … Kessler, R. C. (2011). Cross-national epidemiology of DSM-IV major depressive episode. *BMC Medicine*, *9*, 90. https://doi.org/10.1186/1741-7015-9-90

Bücker, J., Kozicky, J., Torres, I. J., Kauer-Sant'Anna, M., Silveira, L. E., Bond, D. J., Lam, R. W., & Yatham, L. N. (2013). The impact of childhood trauma on cognitive functioning in patients recently recovered from a first manic episode: Data from the Systematic Treatment Optimization Program for Early Mania (STOP-EM). *Journal of Affective Disorders*, *148*(2–3), 424–430. https://doi.org/10.1016/j.jad.2012.11.022

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*, 1–38. https://doi.org/10.1196/annals.1440.011

Bullmore, E. T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., & Brammer, M. J. (1999). Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural mr images of the brain. *IEEE Transactions on Medical Imaging*, *18*(1), 32–42. https://doi.org/10.1109/42.750253

Burdick, K. E., Goldberg, J. F., & Harrow, M. (2010). Neurocognitive dysfunction and psychosocial outcome in patients with bipolar I disorder at 15 year follow-up. *Acta Psychiatrica Scandinavica*, *122*(6), 499–506. https://doi.org/10.1111/j.1600-0447.2010.01590.x

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. https://doi.org/10.1016/j.psychres.2008.03.006

Burdick, K. E., Russo, M., Frangou, S., Mahon, K., Braga, R. J., Shanahan, M., & Malhotra, A. K. (2014). Empirical evidence for discrete neurocognitive subgroups in bipolar disorder: clinical implications. *Psychological Medicine*, *44*(14), 3083–3096. https://doi.org/10.1017/S0033291714000439

Burzynska, A. Z., Wong, C. N., Voss, M. W., Cooke, G. E., McAuley, E., & Kramer, A. F. (2015). White matter integrity supports BOLD signal variability and cognitive performance in the aging human brain. *PLoS ONE*, *10*(4), e0120315. https://doi.org/10.1371/journal.pone.0120315

Butters, M. A., Whyte, E. M., Nebes, R. D., Begley, A. E., Dew, M. A., Mulsant, B. H., Zmuda, M. D., Bhalla, R., Meltzer, C. C., Pollock, B. G., Reynolds, C. F., Becker, J. T., Pollick, B. G., Reynolds, C. F., & Becker, J. T. (2004). The nature and determinants of neuropsychological functioning in late-life depression. *Archives of General Psychiatry*, *61*(6), 587–595. https://doi.org/10.1001/archpsyc.61.6.587

Cambridge Cognition. (2015). *The CANTAB Cognition Handbook*. Cambridge Cognition. www.cantab.com

Camelo, E. V. M., Velasques, B., Ribeiro, P., Netto, T., & Cheniaux, E. (2013). Attention impairment in bipolar disorder: A systematic review. *Psychology and Neuroscience*, *6*(3), 299–309. https://doi.org/10.3922/j.psns.2013.3.08

Cardenas, S. A., Kassem, L., Brotman, M. A., Leibenluft, E., & McMahon, F. J. (2016). Neurocognitive functioning in euthymic patients with bipolar disorder and unaffected relatives: A review of the

literature. *Neuroscience and Biobehavioral Reviews*, *69*, 193–215. https://doi.org/10.1016/j.neubiorev.2016.08.002

Cardoso, T., Bauer, I. E., Meyer, T. D., Kapczinski, F., & Soares, J. C. (2015). Neuroprogression and Cognitive Functioning in Bipolar Disorder: A Systematic Review. *Current Psychiatry Reports*, *17*(75). https://doi.org/10.1007/s11920-015-0605-x

Castaneda, A. E., Suvisaari, J., Marttunen, M., Perälä, J., Saarni, S. I., Aalto-Setälä, T., Aro, H., Koskinen, S., Lönnqvist, J., & Tuulio-Henriksson, A. (2008). Cognitive functioning in a population-based sample of young adults with a history of non-psychotic unipolar depressive disorders without psychiatric comorbidity. *Journal of Affective Disorders*, *110*(1–2), 36–45. https://doi.org/10.1016/j.jad.2007.12.239

Cegalis, J., & Bowlin, J. (1991). *Vigil: Software for the Assessment of Attention*. Forthought.

Cepeda, N. J., Blackwell, K. A., & Munakata, Y. (2013). Speed isn't everything: Complex processing speed measures mask individual differences and developmental changes in executive control. *Developmental Science*, *16*(2), 269–286. https://doi.org/10.1111/desc.12024

Chakrabarty, T., Torres, I. J., Su, W. W., Sawatzky, R., Keramatian, K., & Yatham, L. N. (2021). Cognitive subgroups in first episode bipolar I disorder: Relation to clinical and brain volumetric variables. *Acta Psychiatrica Scandinavica*, *143*(2), 151–161. https://doi.org/10.1111/acps.13245

Chang, C. T., Chang, Y. H., Yung-Wei Wu, J., Lee, S. Y., Chen, S. L., Chen, S. H., Chu, C. H., See Chen, P., Hui Lee, I., Lieh Yeh, T., Tzeng, N. S., Huang, S. Y., Yang, Y. K., Yang, H. F., & Lu, R. B. (2012). Neuropsychological functions impairment in different subtypes of bipolar disorder with or without comorbid anxiety disorders. *Psychiatry Research*, *200*(2–3), 246–251. https://doi.org/10.1016/j.psychres.2012.06.012

Chen, L., Wang, Y., Niu, C., Zhong, S., Hu, H., Chen, P., Zhang, S., Chen, G., Deng, F., Lai, S., Wang, J., Huang, L., & Huang, R. (2018). Common and distinct abnormal frontal-limbic system structural and functional patterns in patients with major depression and bipolar disorder. *NeuroImage: Clinical*, *20*(June), 42–50. https://doi.org/10.1016/j.nicl.2018.07.002

Cheung, E. Y. W., Halari, R., Cheng, K. M., Leung, S. K., & Young, A. H. (2013). Cognitive performance is impaired in euthymic Chinese patients with Bipolar 1 Disorder. *Journal of Affective Disorders*, *151*(1), 156–163. https://doi.org/10.1016/j.jad.2013.05.070

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cohen, R. (2014). *The Neuropsychology of Attention* (2nd ed.). Springer.

Cohen, R. (2019). Continuous Performance Tests. In J. Kreutzer, J. DeLuca, B. Caplan, & SprinkerLink (Eds.), *Encyclopedia of Clinical Neuropsychology* (pp. 699–701). Springer Nature Living Reference.

Cole, M. W., & Schneider, W. (2007). The cognitive control network: Integrated cortical regions with dissociable functions. *NeuroImage*, *37*(1), 343–360. https://doi.org/10.1016/j.neuroimage.2007.03.071

Collantoni, E., Madan, C. R., Meneguzzo, P., Chiappini, I., Tenconi, E., Manara, R., Favaro, A., & Giustiniani, V. (2019). Cortical Complexity in Anorexia Nervosa: A Fractal Dimension Analysis. *Journal of Clincial Medicine*, *9*, 833. https://doi.org/10.3390/jcm9030833

Conners, C. K., Epstein, J. N., Angold, A., & Klaric, J. (2003). Continuous Performance Test Performance in a Normative Epidemiological Sample. *Journal of Abnormal Child Psychology*, *31*(5), 555–562.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed,

and general fluid intelligence. *Intelligence*, *30*(2), 163–183. https://doi.org/10.1016/S0160-2896(01)00096-4

Cotrena, C., Branco, L. D., Shansis, F. M., & Fonseca, R. P. (2016). Executive function impairments in depression and bipolar disorder: Association with functional impairment and quality of life. *Journal of Affective Disorders*, *190*, 744–753. https://doi.org/10.1016/j.jad.2015.11.007

Cotrena, C., Damiani Branco, L., Ponsoni, A., Milman Shansis, F., & Paz Fonseca, R. (2017). Neuropsychological Clustering in Bipolar and Major Depressive Disorder. *Journal of the International Neuropsychological Society*, *23*(7), 584–593. https://doi.org/10.1017/S1355617717000418

Cotrena, C., Damiani Branco, L., Ponsoni, A., Samamé, C., Milman Shansis, F., & Paz Fonseca, R. (2020). Executive functions and memory in bipolar disorders I and II: new insights from meta-analytic results. *Acta Psychiatrica Scandinavica*, *141*(2), 110–130. https://doi.org/10.1111/acps.13121

Cullen, B., Ward, J., Graham, N. A., Deary, I. J., Pell, J. P., Smith, D. J., & Evans, J. J. (2016). Prevalence and correlates of cognitive impairment in euthymic adults with bipolar disorder: A systematic review. *Journal of Affective Disorders*, *205*, 165–181. https://doi.org/10.1016/j.jad.2016.06.063

Daban, C., Mathieu, F., Raust, A., Cochet, B., Scott, J., Etain, B., Leboyer, M., & Bellivier, F. (2012). Is processing speed a valid cognitive endophenotype for bipolar disorder? *Journal of Affective Disorders*, *139*, 98–101. https://doi.org/10.1016/j.jad.2012.02.028

Daniel, B. D., Montali, A., Gerra, M. L., Innamorati, M., Girardi, P., Pompili, M., & Amore, M. (2013). Cognitive impairment and its associations with the path of illness in affective disorders: A comparison between patients with bipolar and unipolar depression in remission. *Journal of Psychiatric Practice*, *19*(4), 275–287. https://doi.org/10.1097/01.pra.0000432597.79019.e2

Dattalo, P. V. (2014). A demonstration of canonical correlation analysis with orthogonal rotation to facilitate interpretation. *VCU Scholars Compass, Social Work Publications*, 1–34. https://scholarscompass.vcu.edu/cgi/viewcontent.cgi?article=1001&context=socialwork_pubs

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience*, *11*(3), 201–211. https://doi.org/10.1038/nrn2793

Demmo, C., Lagerberg, T. V., Aminoff, S. R., Hellvin, T., Kvitland, L. R., Simonsen, C., Andreassen, O. A., Melle, I., & Ueland, T. (2016). History of psychosis and previous episodes as potential explanatory factors for neurocognitive impairment in first-treatment bipolar I disorder. *Bipolar Disorders*, *18*(2), 136–147. https://doi.org/10.1111/bdi.12377

den Hartog, H. M., Derix, M. M. A., Van Bemmel, A. L., Kremer, B., & Jolles, J. (2003). Cognitive functioning in young and middle-aged unmedicated out-patients with major depression: Testing the effort and cognitive speed hypotheses. *Psychological Medicine*, *33*(8), 1443–1451. https://doi.org/10.1017/S003329170300833X

Deng, W., Zhang, B., Zou, W., Zhang, X., Cheng, X., Guan, L., Lin, Y., Lao, G., Ye, B., Li, X., Yang, C., Ning, Y., & Cao, L. (2019). Abnormal degree centrality associated with cognitive dysfunctions in early bipolar disorder. *Frontiers in Psychiatry*, *10*, 140. https://doi.org/10.3389/fpsyt.2019.00140

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Dickersin, K. (1997). How important is publication bias? A synthesis of available data. *AIDS Education and Prevention*, *9*, 15–21.

Dickerson, B. C., Fenstermacher, E., Salat, D. H., Wolk, D. A., Maguire, R. P., Desikan, R., Pacheco, J., Quinn, B. T., Van der Kouwe, A., Greve, D. N., Blacker, D., Albert, M. S., Killiany, R. J., & Fischl, B. (2008). Detection of cortical thickness correlates of cognitive performance: Reliability across MRI scan sessions, scanners, and field strengths. *NeuroImage*, *39*(1), 10–18. https://doi.org/10.1016/j.neuroimage.2007.08.042

Dinga, R., Schmaal, L., Penninx, B. W. J. H., van Tol, M. J., Veltman, D. J., van Velzen, L., Mennes, M., van der Wee, N. J. A., & Marquand, A. F. (2019). Evaluating the evidence for biotypes of depression: Methodological replication and extension of Drysdale et al. (2017). *NeuroImage: Clinical*, *22*, 101796. https://doi.org/10.1016/j.nicl.2019.101796

Doan, N. T., Kaufmann, T., Bettella, F., Jørgensen, K. N., Brandt, C. L., Moberget, T., Alnæs, D., Douaud, G., Duff, E., Djurovic, S., Melle, I., Ueland, T., Agartz, I., Andreassen, O. A., & Westlye, L. T. (2017). Distinct multivariate brain morphological patterns and their added predictive value with cognitive and polygenic risk scores in mental disorders. *NeuroImage: Clinical*, *15*(August 2016), 719–731. https://doi.org/10.1016/j.nicl.2017.06.014

Dobri, M. L., Prince, T., Diaz, A. P., Zunta-soares, G. B., Selvaraj, S., Machado-vieira, R., Meyer, T. D., Sanches, M., & Soares, J. C. (2022). Processing speed — A potential candidate cognitive endophenotype for bipolar disorder. *Journal of Affective Disorders Reports*, *11*, 100459. https://doi.org/10.1016/j.jadr.2022.100459

Donohoe, G., Duignan, A., Hargreaves, A., Morris, D. W., Rose, E., Robertson, D., Cummings, E., Moore, S., Gill, M., & Corvin, A. (2012). Social cognition in bipolar disorder versus schizophrenia: Comparability in mental state decoding deficits. *Bipolar Disorders*, *14*(7), 743–748. https://doi.org/10.1111/bdi.12011

Doose-Grünefeld, S., Eickhoff, S. B., & Müller, V. I. (2015). Audiovisual emotional processing and neurocognitive functioning in patients with depression. *Frontiers in Integrative Neuroscience*, *9*(JAN), 1–13. https://doi.org/10.3389/fnint.2015.00003

Douglas, K. M., Gallagher, P., Robinson, L. J., Carter, J. D., McIntosh, V. V. W., Frampton, C. M. A., Watson, S., Young, A. H., Ferrier, I. N., & Porter, R. J. (2018). Prevalence of cognitive impairment in major depression and bipolar disorder. *Bipolar Disorders*, *20*(3), 260–274. https://doi.org/10.1111/bdi.12602

Douglas, K. M., Porter, R. J., Knight, R. G., & Maruff, P. (2011). Neuropsychological changes and treatment response in severe depression. *British Journal of Psychiatry*, *198*(2), 115–122. https://doi.org/10.1192/bjp.bp.110.080713

Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., Fetcho, R. N., Zebley, B., Oathes, D. J., Etkin, A., Schatzberg, A. F., Sudheimer, K., Keller, J., Mayberg, H. S., Gunning, F. M., Alexopoulos, G. S., Fox, M. D., Pascual-Leone, A., Voss, H. U., … Liston, C. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature Medicine*, *23*(1), 28–38. https://doi.org/10.1038/nm.4246

Duan, Y., Wei, J., Geng, W., Jiang, J., Zhao, X., Li, T., Jiang, Y., Shi, L., Cao, J., Zhu, G., Zhang, K., & Yu, X. (2021). Research on Cognitive Function in Anxious Depression Patients in China. *Journal of Affective Disorders*, *280*(PA), 121–126. https://doi.org/10.1016/j.jad.2020.11.078

Duff, K., Schoenberg, M. R., Scott, J. G., & Adams, R. L. (2005). The relationship between executive functioning and verbal and visual learning and memory. *Archives of Clinical Neuropsychology*, *20*(1), 111–122. https://doi.org/10.1016/j.acn.2004.03.003

Dusi, N., De Carlo, V., Delvecchio, G., Bellani, M., Soares, J. C., & Brambilla, P. (2019). MRI features of clinical outcome in bipolar disorder: A selected review: Special Section on "Translational and Neuroscience Studies in Affective Disorders". Section Editor, Maria Nobile MD, PhD. This Section of JAD focuses on the relevance of translationa. *Journal of Affective Disorders*, *243*, 559–563. https://doi.org/10.1016/j.jad.2018.05.066

Dutta, A., McKie, S., Downey, D., Thomas, E., Juhasz, G., Arnone, D., Elliott, R., Williams, S., Deakin, J. F. W., & Anderson, I. M. (2019). Regional default mode network connectivity in major depressive disorder: modulation by acute intravenous citalopram. *Translational Psychiatry*, *9*(1). https://doi.org/10.1038/s41398-019-0447-0

Eack, S. M., Bahorik, A. L., McKnight, S. A. F., Hogarty, S. S., Greenwald, D. P., Newhill, C. E., Phillips, M. L., Keshavan, M. S., & Minshew, N. J. (2013). Commonalities in social and non-social cognitive impairments in adults with autism spectrum disorder and schizophrenia. *Schizophrenia Research*, *148*(1–3), 24–28. https://doi.org/10.1016/j.schres.2013.05.013

Ebert, B., Miskowiak, K. W., Kloster, M., Johansen, J., Eckholm, C., Wærner, T., Holme, M., & Bruun, L. M. (2017). An ethnographic study of the effects of cognitive symptoms in patients with major depressive disorder: The IMPACT study. *BMC Psychiatry*, *17*(370), 1–13. https://doi.org/10.1186/s12888-017-1523-8

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634.

Ehrminger, M., Brunet-Gouet, E., Cannavo, A.-S. S., Aouizerate, B., Cussac, I., Azorin, J.-M. M., Bellivier, F., Bougerol, T., Courtet, P., Dubertret, C., Etain, B., Kahn, J.-P. P., Leboyer, M., Olié, E., Passerieux, C., & Roux, P. (2019). Longitudinal relationships between cognition and functioning over 2 years in euthymic patients with bipolar disorder: a cross-lagged panel model approach with the FACE-BD cohort. *The British Journal of Psychiatry*, *218*(2), 1–8. https://doi.org/10.1192/bjp.2019.180

Elshahawi, H. H., Essawi, H., Rabie, M. A., Mansour, M., Beshry, Z. A., & Mansour, A. N. (2011). Cognitive functions among euthymic bipolar i patients after a single manic episode versus recurrent episodes. *Journal of Affective Disorders*, *130*(1–2), 180–191. https://doi.org/10.1016/j.jad.2010.10.027

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating Psychological Networks and their Accuracy: A Tutorial Paper. *Behaviour Research Methods*, *50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network Visualizations of Relationships in Psychometric Data. *Journal of Statistical Software*, *48*(4). https://doi.org/10.18637/jss.v048.i04

Epskamp, S., & Fried, E. I. (2018). A Tutorial on Regularized Partial Correlation Networks. *Psychological Methods*, *23*(4), 617–634. https://doi.org/10.1103/PhysRevB.69.161303

Ernst, A. F., & Albers, C. J. (2017). Regression assumptions in clinical psychology research practice-a systematic review of common misconceptions. *PeerJ*, *2017*(5), 1–16. https://doi.org/10.7717/peerj.3323

Erol, A., Kosger, F., Putgul, G., & Ersoy, B. (2014). Ventral prefrontal executive function impairment as a potential endophenotypic marker for bipolar disorder. *Nordic Journal of Psychiatry*, *68*(1), 18–23. https://doi.org/10.3109/08039488.2012.756062

Esan, O., Oladele, O., Adediran, K. I., & Abiona, T. O. (2020). Neurocognitive Impairments (NCI) in bipolar disorder: Comparison with schizophrenia and healthy controls. *Journal of Affective Disorders*, *277*, 175–181. https://doi.org/10.1016/j.jad.2020.08.015

Esterman, M., Noonan, S. K., Rosenberg, M., & Degutis, J. (2013). In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. *Cerebral Cortex*, *23*(11), 2712–2723. https://doi.org/10.1093/cercor/bhs261

Esterman, M., & Rothlein, D. (2019). Models of sustained attention. *Current Opinion in Psychology*, *29*, 174–180. https://doi.org/10.1016/j.copsyc.2019.03.005

Evans, V. C., Iverson, G. L., Yatham, L. N., & Lam, R. W. (2014). The relationship between neurocognitive and psychosocial functioning in major depressive disorder: A systematic review. *Journal of Clinical Psychiatry*, *75*(12), 1359–1370. https://doi.org/10.4088/JCP.13r08939

Favre, P., Pauling, M., Stout, J., Hozer, F., Sarrazin, S., Abé, C., Alda, M., Alloza, C., Alonso-Lana, S., Andreassen, O. A., Baune, B. T., Benedetti, F., Busatto, G. F., Canales-Rodríguez, E. J., Caseras, X., Chaim-Avancini, T. M., Ching, C. R. K., Dannlowski, U., Deppe, M., … Houenou, J. (2019). Widespread white matter microstructural abnormalities in bipolar disorder: evidence from mega- and meta-analyses across 3033 individuals. *Neuropsychopharmacology*, *44*(13), 2285–2293. https://doi.org/10.1038/s41386-019-0485-6

Fears, S. C., Schür, R., Sjouwerman, R., Service, S. K., Araya, C., Araya, X., Bejarano, J., Knowles, E., Gomez-Makhinson, J., Lopez, M. C., Aldana, I., Teshiba, T. M., Abaryan, Z., Al-Sharif, N. B., Navarro, L., Tishler, T. A., Altshuler, L., Bartzokis, G., Escobar, J. I., … Bearden, C. E. (2015). Brain structure-function associations in multi-generational families genetically enriched for bipolar disorder. *Brain*, *138*(7), 2087–2102. https://doi.org/10.1093/brain/awv106

Fears, S. C., Service, S. K., Kremeyer, B., Araya, C., Araya, X., Bejarano, J., Ramirez, M., Castrillón, G., Gomez-Franco, J., Lopez, M. C., Montoya, G., Montoya, P., Aldana, I., Teshiba, T. M., Abaryan, Z., Al-Sharif, N. B., Ericson, M., Jalbrzikowski, M., Luykx, J. J., … Bearden, C. E. (2014). Multisystem component phenotypes of bipolar disorder for genetic investigations of extended pedigrees. *JAMA Psychiatry*, *71*(4), 375–387. https://doi.org/10.1001/jamapsychiatry.2013.4100

Fergusson, D. M., & Woodward, L. J. (2002). Mental health, educational, and social role outcomes of adolescents with depression. *Archives of General Psychiatry*, *59*(3), 225–231. https://doi.org/10.1001/archpsyc.59.3.225

Fernández-Sevillano, J., Alberich, S., Zorrilla, I., González-Ortega, I., López, M. P., Pérez, V., Vieta, E., González-Pinto, A., & Saíz, P. (2021). Cognition in Recent Suicide Attempts: Altered Executive Function. *Frontiers in Psychiatry*, *12*, 701140. https://doi.org/10.3389/fpsyt.2021.701140

Ferrier, I. N., Stanton, B. R., Kelly, T. P., & Scott, J. (1999). Neuropsychological function in euthymic patients with bipolar disorder. *British Journal of Psychiatry*, *175*, 246–251. https://doi.org/10.1192/bjp.175.3.246

Fineberg, N. A., Haddad, P. M., Carpenter, L., Gannon, B., Sharpe, R., Young, A. H., Joyce, E., Rowe, J., Wellsted, D., Nutt, D. J., & Sahakian, B. J. (2013). The size, burden and cost of disorders of the brain in the UK. *Journal of Psychopharmacology*, *27*(9), 761–770. https://doi.org/10.1177/0269881113495118

Finkel, D., Reynolds, C. A., McArdle, J. J., & Pedersen, N. L. (2005). The longitudinal relationship between processing speed and cognitive ability: Genetic and environmental influences. *Behavior Genetics*, *35*(5), 535–549. https://doi.org/10.1007/s10519-005-3281-5

First, M. B., Spitzer, R. L., Williams, J. B. W., & Gibbon, M. (1995). *Structured Clinical Interview for DSM-IV (SCID-I), Research Version*. Biometrics Research.

Fischl, B. (2012). FreeSurfer. *NeuroImage*, *62*(2), 774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021

Fitts, P. M. (1954). The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. *Journal of Experimental Psychology*, *47*(6), 381–391.

Fleck, D. E., Eliassen, J. C., Durling, M., Lamy, M., Adler, C. M., Delbello, M. P., Shear, P. K., Cerullo, M. A., Lee, J. H., & Strakowski, S. M. (2012). Functional MRI of sustained attention in bipolar mania. *Molecular Psychiatry*, *17*(3), 325–336. https://doi.org/10.1038/mp.2010.108

Foland-Ross, L. C., Thompson, P. M., Sugar, C. A., Madsen, S. K., Shen, J. K., Penfold, C., Ahlf, K., Rasser, P. E., Fischer, J., Yang, Y., Townsend, J., Bookheimer, S. Y., & Altshuler, L. L. (2011).

Investigation of cortical thickness abnormalities in lithium-free adults with bipolar I disorder using cortical pattern matching. *American Journal of Psychiatry*, *168*(5), 530–539. https://doi.org/10.1176/appi.ajp.2010.10060896

Formánek, T., Csajbók, Z., Wolfová, K., Kučera, M., Tom, S., Aarsland, D., & Cermakova, P. (2020). Trajectories of depressive symptoms and associated patterns of cognitive decline. *Scientific Reports*, *10*, 20888. https://doi.org/10.1038/s41598-020-77866-6

Fortenbaugh, F. C., Degutis, J., & Esterman, M. (2017). Recent theoretical, neural, and clinical advances in sustained attention research. *Annals of the New York Academy of Sciences*, *1396*, 70–91. https://doi.org/10.1111/nyas.13318

Frajo-Apor, B., Kemmler, G., Pardeller, S., Huber, M., Macina, C., Welte, A.-S., Hoertnagl, C., & Hofer, A. (2020). Emotional intelligence in bipolar-I-disorder: A comparison between patients, unaffected siblings, and control subjects. *European Psychiatry*, *63*(1), e69, 1–7. https://doi.org/10.1192/j.eurpsy.2020.66

Frangou, S., Kington, J., Raymont, V., & Shergill, S. S. (2008). Examining ventral and dorsal prefrontal function in bipolar disorder: A functional magnetic resonance imaging study. *European Psychiatry*, *23*(4), 300–308. https://doi.org/10.1016/j.eurpsy.2007.05.002

Friston, K. J., Frith, C. D., Frackowiak, R. S. J. J., & Turner, R. (1995). Characterizing dynamic brain responses with fMRI: A multivariate approach. *NeuroImage*, *2*(2), 166–172. https://doi.org/10.1006/nimg.1995.1019

Frydecka, D., Eissa, A. M., Hewedi, D. H., Ali, M., Drapała, J., Misiak, B., Kłosińska, E., Phillips, J. R., & Moustafa, A. A. (2014). Impairments of working memory in schizophrenia and bipolar disorder: The effect of history of psychotic symptoms and different aspects of cognitive task demands. *Frontiers in Behavioral Neuroscience*, *8*(NOV), 1–11. https://doi.org/10.3389/fnbeh.2014.00416

Galimberti, C., Bosi, M. F., Caricasole, V., Zanello, R., Dell'Osso, B., & Viganò, C. A. (2020). Using network analysis to explore cognitive domains in patients with unipolar versus bipolar depression: A prospective naturalistic study. *CNS Spectrums*, *25*, 380–391. https://doi.org/10.1017/S1092852919000968

Gallagher, P. (2021). Neuropsychology of bipolar disorder. In *Current Topics in Behavioral Neurosciences* (pp. 239–253). Springer. https://doi.org/10.1007/7854_2020_148

Gallagher, P., Gray, J. M., & Kessels, R. P. C. (2015). Fractionation of visuo-spatial memory processes in bipolar depression: A cognitive scaffolding account. *Psychological Medicine*, *45*(3), 545–558. https://doi.org/10.1017/S0033291714001676

Gallagher, P., Gray, J. M., Watson, S., Young, A. H., & Ferrier, I. N. (2014). Neurocognitive functioning in bipolar depression: A component structure analysis. *Psychological Medicine*, *44*(5), 961–974. https://doi.org/10.1017/S0033291713001487

Gallagher, P., Nilsson, J., Finkelmeyer, A., Goshawk, M., Macritchie, K. A., Lloyd, A. J., Thompson, J. M., Porter, R. J., Young, A. H., Ferrier, I. N., McAllister-Williams, R. H., & Watson, S. (2015). Neurocognitive intra-individual variability in mood disorders: Effects on attentional response time distributions. *Psychological Medicine*, *45*(14), 2985–2997. https://doi.org/10.1017/S0033291715000926

Gautam, P., Anstey, K. J., Wen, W., Sachdev, P. S., & Cherbuin, N. (2015). Cortical gyrification and its relationships with cortical volume, cortical thickness, and cognitive performance in healthy mid-life adults. *Behavioural Brain Research*, *287*, 331–339. https://doi.org/10.1016/j.bbr.2015.03.018

Genon, S., Eickhoff, S. B., & Kharabian, S. (2022). Linking interindividual variability in brain structure to behaviour. *Nature Reviews Neuroscience*, *23*(5), 307–318. https://doi.org/10.1038/s41583-022-00584-7

Geraets, A. F. J., Schram, M. T., Jansen, J. F. A., Koster, A., Dagnelie, P. C., Van Greevenbroek, M. M. J., Stehouwer, C. D. A., Verhey, F. R. J., & Köhler, S. (2021). The relation of depression with structural brain abnormalities and cognitive functioning: The Maastricht study. *Psychological Medicine*. https://doi.org/10.1017/S0033291721000222

Gilsoul, J., Simon, J., Hogge, M., & Collette, F. (2019). Do attentional capacities and processing speed mediate the effect of age on executive functioning? *Aging, Neuropsychology, and Cognition*, *26*(2), 282–317. https://doi.org/10.1080/13825585.2018.1432746

Gogos, A., Joshua, N., & Rossell, S. L. (2010). Use of the Repeatable Battery for the Assessment of Neuropsychological status (RBANS) to investigate group and gender differences in schizophrenia and bipolar disorder. *Australian and New Zealand Journal of Psychiatry*, *44*(3), 220–229. https://doi.org/10.3109/00048670903446882

Goltermann, J., Redlich, R., Grotegerd, D., Dohm, K., Leehr, E. J., Böhnlein, J., Förster, K., Meinert, S., Enneking, V., Richter, M., Repple, J., DeVillers, I., Kloecker, M., Jansen, A., Krug, A., Nenadić, I., Brosch, K., Meller, T., Stein, F., … Dannlowski, U. (2021). Childhood maltreatment and cognitive functioning: the role of depression, parental education, and polygenic predisposition. *Neuropsychopharmacology*, *46*(5), 891–899. https://doi.org/10.1038/s41386-020-00794-6

Gómez-Benito, J., Guilera, G., Pino, O., Tabarés-Seisdedos, R., & Martínez-Arán, A. (2014). Comparing neurocognitive impairment in schizophrenia and bipolar I disorder using the Screen for Cognitive Impairment in Psychiatry Scale. *International Journal of Clinical and Health Psychology*, *14*(2), 128–136. https://doi.org/10.1016/S1697-2600(14)70046-8

Gorenstein, C., De Carvalho, S. C., Artes, R., Moreno, R. A., & Marcourakis, T. (2006). Cognitive performance in depressed patients after chronic use of antidepressants. *Psychopharmacology*, *185*(1), 84–92. https://doi.org/10.1007/s00213-005-0274-2

Gorwood, P., Corruble, E., Falissard, B., & Goodwin, G. M. (2008). Toxic effects of depression on brain function: Impairment of delayed recall and the cumulative length of depressive disorder in a large sample of depressed outpatients. *American Journal of Psychiatry*, *165*(6), 731–739. https://doi.org/10.1176/appi.ajp.2008.07040574

Goswami, U., Sharma, A., Khastigir, U., Nicol, I. N., Young, A. H., Gallagher, P., Thompson, J. M., & Moore, P. B. (2006). Neuropsychological dysfunction, soft neurological signs and social disability in euthymic patients with bipolar disorder. *British Journal of Psychiatry*, *188*, 366–373.

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. https://doi.org/10.1111/j.1600-0447.2009.01390.x

Grande, I., Berk, M., Birmaher, B., & Vieta, E. (2016). Bipolar disorder. *The Lancet*, *387*(10027), 1561–1572. https://doi.org/10.1016/S0140-6736(15)00241-X

Gruber, M., Klein, H., Mauritz, M., Lange, S. C. De, Grumbach, P., Borgers, T., Enneking, V., Klug, M., Stein, F., & Brosch, K. (2022). The Impact of Cognitive Reserve on Cognitive Impairment, Brain Network Pathology and Disease Trajectory in Major Depressive Disorder. *PsyArXiv Preprint*. https://doi.org/10.31234/osf.io/mxbku

Gu, C. Z., He, H. L., Duan, H. F., Su, Z. H., Chen, H., & Gan, J. L. (2016). Predictors of neurocognitive impairment at 2 years after a first-episode major depressive disorder. *Comprehensive Psychiatry*, *68*, 24–33. https://doi.org/10.1016/j.comppsych.2016.03.009

Gualtieri, C. T., & Johnson, L. G. (2006). Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs. *Archives of Clinical Neuropsychology*, *21*(7), 623–643. https://doi.org/10.1016/j.acn.2006.05.007

Gualtieri, C. T., & Johnson, L. G. (2008). Age-related cognitive decline in patients with mood

disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *32*(4), 962–967. https://doi.org/10.1016/j.pnpbp.2007.12.030

Gunstad, J., Cohen, R. A., Paul, R. H., & Gordon, E. (2006). Dissociation of the component processes of attention in healthy adults. *Archives of Clinical Neuropsychology*, *21*(7), 645–650. https://doi.org/10.1016/j.acn.2006.05.009

Gutiérrez-Galve, L., Bruno, S., Wheeler-Kingshott, C. A. M., Summers, M., Cipolotti, L., & Ron, M. A. (2012). IQ and the fronto-temporal cortex in bipolar disorder. *Journal of the International Neuropsychological Society*, *18*(2), 370–374. https://doi.org/10.1017/S1355617711001706

Halappa, N., Thirthalli, J., Varambally, S., Rao, M., Christopher, R., & Nanjundaiah, G. (2018). Improvement in neurocognitive functions and serum brain-derived neurotrophic factor levels in patients with depression treated with antidepressants and yoga. *Indian Journal of Psychiatry*, *60*(1), 32–37. https://doi.org/10.4103/psychiatry.IndianJPsychiatry_154_17

Halvorsen, M., Hifdt, R. S., Myrbakk, I. N., Wang, C. E. A., Sundet, K., Eisemann, M., & Waterloo, K. (2012). Cognitive function in unipolar major depression: A comparison of currently depressed, previously depressed, and never depressed individuals. *Journal of Clinical and Experimental Neuropsychology*, *34*(7), 782–790. https://doi.org/10.1080/13803395.2012.683853

Hamilton, M. (1960). A rating scale for depression. *J Neurol Neurosurg Psychiatry*, *23*, 56–62.

Hanford, L. C., Nazarov, A., Hall, G. B., & Sassi, R. B. (2016). Cortical thickness in bipolar disorder: A systematic review. *Bipolar Disorders*, *18*(1), 4–18. https://doi.org/10.1111/bdi.12362

Harmer, C. J., Clark, L., Grayson, L., & Goodwin, G. M. (2002). Sustained attention deficit in bipolar disorder is not a working memory impairment in disguise. *Neuropsychologia*, *40*(9), 1586–1590. https://doi.org/10.1016/S0028-3932(02)00019-2

Haro, J. M., Hammer-Helmich, L., Saragoussi, D., Ettrup, A., & Larsen, K. G. (2019). Patient-reported depression severity and cognitive symptoms as determinants of functioning in patients with major depressive disorder : a secondary analysis of the 2-year prospective PERFORM study. *Neuropsychiatric Disease and Treatment*, *15*, 2313–2323. https://doi.org/10.2147/NDT.S206825

Harrell, F. E. (2016). *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis* (2nd ed.). Springer Cham. https://doi.org/10.1007/978-3-319-19425-7

Hartberg, C. B., Sundet, K., Rimol, L. M., Haukvik, U. K., Lange, E. H., Nesvåg, R., Dale, A. M., Melle, I., Andreassen, O. A., & Agartz, I. (2011). Brain cortical thickness and surface area correlates of neurocognitive performance in patients with schizophrenia, bipolar disorder, and healthy adults. *Journal of the International Neuropsychological Society*, *17*(6), 1080–1093. https://doi.org/10.1017/S1355617711001081

Hartberg, C. B., Sundet, K., Rimol, L. M., Haukvik, U. K., Lange, E. H., Nesvåg, R., Melle, I., Andreassen, O. A., & Agartz, I. (2011). Subcortical brain volumes relate to neurocognition in schizophrenia and bipolar disorder and healthy controls. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *35*(4), 1122–1130. https://doi.org/10.1016/j.pnpbp.2011.03.014

Hatton, S. N., Lagopoulos, J., Hermens, D. F., Scott, E., Hickie, I. B., & Bennett, M. R. (2013). Cortical thinning in young psychosis and bipolar patients correlate with common neurocognitive deficits. *International Journal of Bipolar Disorders*, *1*(1), 1–13. https://doi.org/10.1186/2194-7511-1-3

Henry, J. D., & Crawford, J. R. (2005). A meta-analytic review of verbal fluency deficits in depression. *Journal of Clinical and Experimental Neuropsychology*, *27*(1), 78–101. https://doi.org/10.1080/138033990513654

Hibar, D. P., Westlye, L. T., Doan, N. T., Jahanshad, N., Cheung, J. W., Ching, C. R. K., Versace, A.,

Bilderbeck, A. C., Uhlmann, A., Mwangi, B., Krämer, B., Overs, B., Hartberg, C. B., Abe, C., Dima, D., Grotegerd, D., Sprooten, E., Ben, E., Jimenez, E., … Andreassen, O. A. (2018). Cortical abnormalities in bipolar disorder: An MRI analysis of 6503 individuals from the ENIGMA Bipolar Disorder Working Group. *Molecular Psychiatry*, *23*(4), 932–942. https://doi.org/10.1038/mp.2017.73

Hibar, D. P., Westlye, L. T., Van Erp, T. G. M., Rasmussen, J., Leonardo, C. D., Faskowitz, J., Haukvik, U. K., Hartberg, C. B., Doan, N. T., Agartz, I., Dale, A. M., Gruber, O., Krämer, B., Trost, S., Liberg, B., Abé, C., Ekman, C. J., Ingvar, M., Landén, M., … Andreassen, O. A. (2016). Subcortical volumetric abnormalities in bipolar disorder. *Molecular Psychiatry*, *21*(12), 1710–1716. https://doi.org/10.1038/mp.2015.227

Higgins, J. P. T., Li, T., & Deeks, J. J. (2022). Chapter 6: Choosing effect measures and computing estimates of effect. In J. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. Page, & V. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions version 6.3*. Cochrane. https://doi.org/ch6

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M., & Welch, V. (2022). Cochrane Handbook for Systematic Reviews of Interventions. In *Cochrane Handbook for Systematic Reviews of Interventions version 6.3*. Cochrane. www.training.cochrane.org/handbook

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hill, S. K., Reilly, J. L., Harris, M. S. H., Rosen, C., Marvin, R. W., DeLeon, O., & Sweeney, J. A. (2009). A comparison of neuropsychological dysfunction in first-episode psychosis patients with unipolar depression, bipolar disorder, and schizophrenia. *Schizophrenia Research*, *113*(2), 167–175. https://doi.org/10.1016/j.schres.2009.04.020

Hirschfeld, M., & Vornik, L. (2005). Bipolar Disorder - Costs and Comorbidity. *The American Journal of Managed Care*, *11*(3), S85–S90.

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., & Zarate, C. A. (2008). A comparison of cognitive functioning in medicated and unmedicated subjects with bipolar depression. *Bipolar Disorders*, *10*(7), 806–815. https://doi.org/10.1111/j.1399-5618.2008.00628.x

Hou, Y., Yao, S., Hu, S., Zhou, Q., Han, H., Yu, X., McIntyre, R. S., & Shi, C. (2020). PSYCHOMETRIC properties of the Chinese version of the THINC-it tool for cognitive symptoms in patients with major depressive disorder. *Journal of Affective Disorders*, *273*, 586–591. https://doi.org/10.1016/j.jad.2020.03.146

Hsu, S. E., Chin Chen, K., Lee, L. T., Chun Tsai, H., Lee, I. H., See Chen, P., & Yang, Y. K. (2015). Comparison of cognitive deficits among drug-naive patients with schizophrenia and major depressive disorder. *Journal of Affective Disorders*, *175*, 133–138. https://doi.org/10.1016/j.jad.2014.12.059

Hu, Y., Li, J., Zhao, Y., Dong, Z., Qiu, P., Yang, S., Xu, H., Kuang, W., & Gong, Q. (2022). Memory and processing speed impairments in first-episode drug- naïve patients with major depressive disorder. *Journal of Affective Disorders*. https://doi.org/10.1016/j.jad.2022.10.048

Humes, G. E., Welsh, M. G., Retzlaff, P., & Cookson, N. (1997). Towers of Hanoi and London: Reliability and validity of two executive function tasks. *Assessment*, *4*(3), 249–257. https://doi.org/10.1177/107319119700400305

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology*, *45*(3), 188–196. https://doi.org/10.1037/h0056940

Iverson, G. L., Brooks, B. L., Langenecker, S. A., & Young, A. H. (2011). Identifying a cognitive

impairment subgroup in adults with mood disorders. *Journal of Affective Disorders*, *132*(3), 360–367. https://doi.org/10.1016/j.jad.2011.03.001

Jakobsen, L. H., Sorensen, J. M., Rask, I. K., Jensen, B., & Kondrup, J. (2011). Validation of reaction time as a measure of cognitive function and quality of life in healthy subjects and patients. *Nutrition*, *27*(5), 561–570. https://doi.org/10.1016/j.nut.2010.08.003

Jamieson, A., Goodwill, A. M., Termine, M., Campbell, S., & Szoeke, C. (2019). Depression related cerebral pathology and its relationship with cognitive functioning: A systematic review. *Journal of Affective Disorders*, *250*, 410–418. https://doi.org/10.1016/j.jad.2019.03.042

Jensen, A., & Munro, E. (1979). Reaction time, movement time, and intelligence. *Intelligence*, *3*(2), 121–126.

Jensen, J., Knorr, U., Vinberg, M., Kessing, L. V., & Miskowiak, K. W. (2016). Discrete neurocognitive subgroups in fully or partially remitted bipolar disorder: Associations with functional abilities. *Journal of Affective Disorders*, *205*, 378–386. https://doi.org/10.1016/j.jad.2016.08.018

Ji, Y., Li, W., Liu, B., Liu, J., Ju, Y., Wang, M., Chen, Y., & Li, L. (2020). Clinical characteristics of cognitive deficits in major depressive disorder: A 6-month prospective study. *Revista de Psiquiatria Clinica*, *47*(4), 101–105. https://doi.org/10.1590/0101-60830000000241

Jiménez-López, E., Aparicio, A. I., Sánchez-Morla, E. M., Rodriguez-Jimenez, R., Vieta, E., & Santos, J. L. (2017). Neurocognition in patients with psychotic and non-psychotic bipolar I disorder. A comparative study with individuals with schizophrenia. *Journal of Affective Disorders*, *222*(July), 169–176. https://doi.org/10.1016/j.jad.2017.07.014

Jin, K., Lu, J., Yu, Z., Shen, Z., Li, H., Mou, T., Xu, Y., & Huang, M. (2020). Linking peripheral IL-6, IL-1β and hypocretin-1 with cognitive impairment from major depression. *Journal of Affective Disorders*, *277*, 204–211. https://doi.org/10.1016/j.jad.2020.08.024

Jørgensen, J. L., Macoveanu, J., Petersen, J. Z., Knudsen, G. M., Kessing, L. V., Jørgensen, M. B., & Miskowiak, K. W. (2023). Association of childhood trauma with cognitive impairment and structural brain alterations in remitted patients with bipolar disorder. *Journal of Affective Disorders*. https://doi.org/10.1016/J.JAD.2023.05.078

Joshi, G., Biederman, J., Petty, C., Goldin, R. L., Furtak, S. L., & Wozniak, J. (2013). Examining the Comorbidity of Bipolar Disorder and Autism Spectrum Disorders: A Large Controlled Analysis of Phenotypic and Familial Correlates in a Referred Population of Youth With Bipolar I Disorder With and Without Autism Spectrum Disorders. *The Journal of Clinical Psychiatry*, *74*(6), 6865. https://doi.org/10.4088/JCP.12m07392

Joy, S., Fein, D., & Kaplan, E. (2003). Decoding Digit Symbol: Speed, Memory, and Visual Scanning. *Assessment*, *10*(1), 56–65. https://doi.org/10.1177/0095399702250335

Joy, S., Fein, D., Kaplan, E., & Freedman, M. (2000). Speed and memory in WAIS-R-NI Digit Symbol performance among healthy older adults. *Journal of the International Neuropsychological Society*, *6*(7), 770–780. https://doi.org/10.1017/S1355617700677044

Joy, S., Kaplan, E., & Fein, D. (2003). Digit Symbol - Incidental Learning in the WAIS-III: Construct Validity and Clinical Significance. *Clinical Neuropsychologist*, *17*(2), 182–194. https://doi.org/10.1076/clin.17.2.182.16495

Joy, S., Kaplan, E., & Fein, D. (2004). Speed and memory in the WAIS-III Digit Symbol - Coding subtest across the adult lifespan. *Archives of Clinical Neuropsychology*, *19*(6), 759–767. https://doi.org/10.1016/j.acn.2003.09.009

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica*, *86*, 199–225.

Kang, Y., Kang, W., Han, K. M., Tae, W. S., & Ham, B. J. (2022). Associations between cognitive

impairment and cortical thickness alterations in patients with euthymic and depressive bipolar disorder. *Psychiatry Research: Neuroimaging*, *322*, 111462. https://doi.org/10.1016/j.pscychresns.2022.111462

Karantonis, J. A., Carruthers, S. P., Burdick, K. E., Pantelis, C., Green, M., Rossell, S. L., Hughes, M. E., Cropley, V., & Van Rheenen, T. E. (2023). Brain Morphological Characteristics of Cognitive Subgroups of Schizophrenia-Spectrum Disorders and Bipolar Disorder: A Systematic Review with Narrative Synthesis. *Neuropsychology Review*, *33*, 192–220. https://doi.org/10.1007/s11065-021-09533-0

Karantonis, J. A., Carruthers, S. P., Rossell, S. L., Pantelis, C., Hughes, M., Wannan, C., Cropley, V., & Van Rheenen, T. E. (2021). A Systematic Review of Cognition-Brain Morphology Relationships on the Schizophrenia-Bipolar Disorder Spectrum. *Schizophrenia Bulletin*, *47*(6), 1557–1600. https://doi.org/10.1093/schbul/sbab054

Karantonis, J. A., Rossell, S. L., Carruthers, S. P., Sumner, P., Hughes, M., Green, M. J., Pantelis, C., Burdick, K. E., Cropley, V., & Van Rheenen, T. E. (2020). Cognitive validation of cross-diagnostic cognitive subgroups on the schizophrenia-bipolar spectrum. *Journal of Affective Disorders*, *266*, 710–721. https://doi.org/10.1016/j.jad.2020.01.123

Karantonis, J. A., Rossell, S. L., Carruthers, S. P., Sumner, P., Hughes, M., Green, M. J., Pantelis, C., Burdick, K. E., Cropley, V., & Van Rheenen, T. E. (2021). Brain morphology does not clearly map to cognition in individuals on the bipolar-schizophrenia-spectrum: a cross-diagnostic study of cognitive subgroups. *Journal of Affective Disorders*, *281*, 776–785. https://doi.org/10.1016/j.jad.2020.11.064

Kebets, V., Holmes, A. J., Orban, C., Tang, S., Li, J., Sun, N., Kong, R., Poldrack, R. A., & Yeo, B. T. T. (2019). Somatosensory-Motor Dysconnectivity Spans Multiple Transdiagnostic Dimensions of Psychopathology. *Biological Psychiatry*, *86*(10), 779–791. https://doi.org/10.1016/j.biopsych.2019.06.013

Kempton, M. J., Geddes, J. R., Ettinger, U., Williams, S. C. R., & Grasby, P. M. (2008). Meta-analysis, database, and meta-regression of 98 structural imaging studies in bipolar disorder. *Archives of General Psychiatry*, *65*(9), 1017–1032. https://doi.org/10.1001/archpsyc.65.9.1017

Kertzman, S., Reznik, I., Hornik-Lurie, T., Weizman, A., Kotler, M., & Amital, D. (2010). Stroop performance in major depression: Selective attention impairment or psychomotor slowness? *Journal of Affective Disorders*, *122*(1–2), 167–173. https://doi.org/10.1016/j.jad.2009.08.009

Kieseppä, T., Mäntylä, R., Tuulio-Henriksson, A., Luoma, K., Mantere, O., Ketokivi, M., Holma, M., Jylhä, P., Melartin, T., Suominen, K., Vuorilehto, M., & Isometsä, E. (2014). White matter hyperintensities and cognitive performance in adult patients with bipolar I, bipolar II, and major depressive disorders. *European Psychiatry*, *29*(4), 226–232. https://doi.org/10.1016/j.eurpsy.2013.08.002

Kieseppä, T., Tuulio-Henriksson, A., Haukka, J., Van Erp, T., Glahn, D., Cannon, T. D., Partonen, T., Kaprio, J., & Lönnqvist, J. (2005). Memory and verbal learning functions in twins with bipolar-I disorder, and the role of information-processing speed. *Psychological Medicine*, *35*(2), 205–215. https://doi.org/10.1017/S0033291704003125

Kim, D., Kim, J., Koo, T., Yun, H., & Won, S. (2015). Shared and distinct neurocognitive endophenotypes of schizophrenia and psychotic bipolar disorder. *Clinical Psychopharmacology and Neuroscience*, *13*(1), 94–102. https://doi.org/10.9758/cpn.2015.13.1.94

Kim, G. H., Kim, J. Y., Kim, J. E., Ma, J., Kim, B. R., Im, J. J., Jeong, H. S., Namgung, E., Lee, S., Kang, I., Lyoo, I. K., Hwang, J., & Yoon, S. (2019). Alterations in structural rich-club connectivity of the precuneus are associated with depressive symptoms among individuals with subjective memory complaints. *Cognitive, Affective and Behavioral Neuroscience*, *19*, 73–87. https://doi.org/10.3758/s13415-018-0645-x

Kim, J. U., Bessette, K. L., Westlund-Schreiner, M., Pocius, S., Dillahunt, A. K., Frandsen, S., Thomas, L., Easter, R., Skerrett, K., Stange, J. P., Welsh, R. C., Langenecker, S. A., & Koppelmans, V. (2022). Relations of gray matter volume to dimensional measures of cognition and affect in mood disorders. *Cortex*, 100001. https://doi.org/10.1016/j.cortex.2022.06.019

Kim, W. J., Ha, R. Y., Sun, J. Y., Ryu, V., Lee, S. J., Ha, K., & Cho, H. S. (2014). Autobiographical memory and its association with neuropsychological function in bipolar disorder. *Comprehensive Psychiatry*, *55*(2), 290–297. https://doi.org/10.1016/j.comppsych.2013.09.010

King, J. B., & Anderson, J. S. (2018). Sustained versus instantaneous connectivity differentiates cognitive functions of processing speed and episodic memory. *Human Brain Mapping*, *39*(12), 4949–4961. https://doi.org/10.1002/hbm.24336

Kjærstad, H. L., Eikeseth, F. F., Vinberg, M., Kessing, L. V., & Miskowiak, K. W. (2021). Neurocognitive heterogeneity in patients with bipolar disorder and their unaffected relatives: Associations with emotional cognition. *Psychological Medicine*, *51*(4), 668–679. https://doi.org/10.1017/S0033291719003738

Kjaerstad, H. L., Søhol, K., Vinberg, M., Kessing, L. V., & Miskowiak, K. W. (2023). The trajectory of emotional and non-emotional cognitive function in newly diagnosed patients with bipolar disorder and their unaffected relatives: A 16-month follow-up study. *European Neuropsychopharmacology*, *67*, 4–21. https://doi.org/10.1016/j.euroneuro.2022.11.004

Knöchel, C., Reuter, J., Reinke, B., Stäblein, M., Marbach, K., Feddern, R., Kuhlmann, K., Alves, G., Prvulovic, D., Wenzler, S., Linden, D. E. J., & Oertel-Knöchel, V. (2016). Cortical thinning in bipolar disorder and schizophrenia. *Schizophrenia Research*, *172*(1–3), 78–85. https://doi.org/10.1016/j.schres.2016.02.007

Koopowitz, S. M., Maré, K. T., Zar, H. J., Stein, D. J., & Ipser, J. C. (2021). The neurocognitive profile of post-traumatic stress disorder (PTSD), major depressive disorder (MDD), and PTSD with comorbid MDD. *Brain and Behavior*, *11*(4), 1–10. https://doi.org/10.1002/brb3.1950

Krabbendam, L., Honig, A., Wiersma, J., Vuurman, E. F. P. M., Hofman, P. A. M., Derix, M. M. A., Nolen, W. A., & Jolles, J. (2000). Cognitive dysfunctions and white matter lesions in patients with bipolar disorder in remission. *Acta Psychiatrica Scandinavica*, *101*(4), 274–280. https://doi.org/10.1034/j.1600-0447.2000.101004274.x

Kriesche, D., Woll, C. F. J., Tschentscher, N., Engel, R. R., & Karch, S. (2022). Neurocognitive deficits in depression: a systematic review of cognitive impairment in the acute and remitted state. *European Archives of Psychiatry and Clinical Neuroscience*, *0123456789*. https://doi.org/10.1007/s00406-022-01479-5

Kuang, L., Gao, W., Long, Z., Cao, W., Cui, D., Guo, Y., Jiao, Q., Qiu, J., Su, L., & Lu, G. (2022). Common and Specific Characteristics of Adolescent Bipolar Disorder Types I and II: A Combined Cortical Thickness and Structural Covariance Analysis. *Frontiers in Psychiatry*, *12*, 750798. https://doi.org/10.3389/fpsyt.2021.750798

Kurtz, M. M., & Gerraty, R. T. (2009). A Meta-analytic Investigation of Neurocognitive Deficits in Bipolar Illness: Profile and Effects of Clinical State. *Neuropsychology*, *23*(5), 551–562. https://doi.org/10.1037/a0016277

Lambert, Z. V., Wildt, A. R., & Durand, R. M. (1988). Redundancy Analysis: An Alternative to Canonical Correlation and Multivariate Multiple Regression in Exploring Interset Associations. *Psychological Bulletin*, *104*(2), 282–289. https://doi.org/10.1037//0033-2909.104.2.282

Lan, M. J., Chhetry, B. T., Oquendo, M. A., Sublette, M. E., Sullivan, G., Mann, J. J., & Parsey, R. V. (2014). Cortical thickness differences between bipolar depression and major depressive disorder. *Bipolar Disorders*, *16*(4), 378–388. https://doi.org/10.1111/bdi.12175

Lange, R. T., & Lippa, S. M. (2016). Processing Speed Index. In J. Kreutzer, J. DeLuca, & B. Caplan

(Eds.), *Encyclopedia of Clinical Neuropsychology*. Springer, Cham. https://doi.org/10.1007/978-3-319-56782-2

Langenecker, S. A., Saunders, E. F. H., Kade, A. M., Ransom, M. T., & McInnis, M. G. (2010). Intermediate: Cognitive phenotypes in bipolar disorder. *Journal of Affective Disorders*, *122*(3), 285–293. https://doi.org/10.1016/j.jad.2009.08.018

Leclaire, K. N., Osmon, D. C., & Driscoll, I. (2020). A distributional and theoretical analysis of reaction time in the reversal task across adulthood. *Journal of Clinical and Experimental Neuropsychology*, *42*(2), 199–207. https://doi.org/10.1080/13803395.2019.1703909

Lee, R., Hermens, D. F., Naismith, S. L., Lagopoulos, J., Jones, A., Scott, J., Chitty, K. M., White, D., Robillard, R., Scott, E. M., & Hickie, I. B. (2015). Neuropsychological and functional outcomes in recent-onset major depression, bipolar disorder and schizophrenia-spectrum disorders: A longitudinal cohort study. *Translational Psychiatry*, *5*(4), e555-10. https://doi.org/10.1038/tp.2015.50

Lee, R., Hermens, D. F., Scott, J., Redoblado-Hodge, M. A., Naismith, S. L., Lagopoulos, J., Griffiths, K. R., Porter, M. A., & Hickie, I. B. (2014). A meta-analysis of neuropsychological functioning in first-episode bipolar disorders. *Journal of Psychiatric Research*, *57*(1), 1–11. https://doi.org/10.1016/j.jpsychires.2014.06.019

Lee, S. Y., Wang, L. J., Chang, C. H., Wu, C. C., Chen, H. L., Lin, S. H., Chu, C. L., Lu, T., & Lu, R. B. (2017). Serum DHEA-S concentration correlates with clinical symptoms and neurocognitive function in patients with bipolar II disorder: A case-controlled study. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *74*, 31–35. https://doi.org/10.1016/j.pnpbp.2016.11.006

Leech, R., Kamourieh, S., Beckmann, C. F., & Sharp, D. J. (2011). Fractionating the Default Mode Network: Distinct Contributions of the Ventral and Dorsal Posterior Cingulate Cortex to Cognitive Control. *Journal of Neuroscience*, *31*(9), 3217–3224. https://doi.org/10.1523/jneurosci.5626-10.2011

Leech, R., & Sharp, D. J. (2014). The role of the posterior cingulate cortex in cognition and disease. *Brain*, *137*(1), 12–32. https://doi.org/10.1093/brain/awt162

Lemke, H., Klute, H., Skupski, J., Thiel, K., Waltemate, L., Winter, A., Breuer, F., Meinert, S., Klug, M., Enneking, V., Winter, N. R., Grotegerd, D., Leehr, E. J., Repple, J., Dohm, K., Opel, N., Stein, F., Meller, T., Brosch, K., … Dannlowski, U. (2022). Brain structural correlates of recurrence following the first episode in patients with major depressive disorder. *Translational Psychiatry*, *12*(1), 349. https://doi.org/10.1038/s41398-022-02113-7

Leung, M. M. W., Lui, S. S. Y., Wang, Y., Tsui, C. F., Au, A. C. W., Yeung, H. K. H., Yang, T. xiao, Li, Z., Cheng, C. wai, Cheung, E. F. C., & Chan, R. C. K. (2016). Patients with bipolar disorder show differential executive dysfunctions: A case-control study. *Psychiatry Research*, *238*, 129–136. https://doi.org/10.1016/j.psychres.2016.01.055

Levada, O. A., & Troyan, A. S. (2019). Cognitive-functional relationships in major depressive disorder: Crucial data from a Ukrainian open-label study of vortioxetine versus escitalopram. *Journal of Affective Disorders*, *250*(February), 114–122. https://doi.org/10.1016/j.jad.2019.03.040

Lewandowski, K. E., Baker, J. T., McCarthy, J. M., Norris, L. A., & Öngür, D. (2018). Reproducibility of cognitive profiles in psychosis using cluster analysis. *Journal of the International Neuropsychological Society*, *24*(4), 382–390. https://doi.org/10.1017/S1355617717001047

Lewandowski, K. E., Cohen, T. R., & Ongur, D. (2020). Cognitive and clinical predictors of community functioning across the psychoses. *PsyCh Journal*, *9*(2), 163–173. https://doi.org/10.1002/pchj.356

Lewandowski, K. E., Whitton, A. E., Pizzagalli, D. A., Norris, L. A., Ongur, D., & Hall, M. H. (2016).

Reward learning, neurocognition, social cognition, and symptomatology in psychosis. *Frontiers in Psychiatry*, *7*(JUN). https://doi.org/10.3389/fpsyt.2016.00100

Leyfer, O. T., Folstein, S. E., Bacalman, S., Davis, N. O., Dinh, E., Morgan, J., Tager-Flusberg, H., & Lainhart, J. E. (2006). Comorbid psychiatric disorders in children with autism: Interview development and rates of disorders. *Journal of Autism and Developmental Disorders*, *36*(7), 849–861. https://doi.org/10.1007/s10803-006-0123-0

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). Oxford University Press.

Lezak, M. D., Howieson, D. B., Loring, D. W., & Fischer, J. S. (2004). *Neuropsychological Assessment*. Oxford University Press.

Li, X., Morgan, P. S., Ashburner, J., Smith, J., & Rorden, C. (2016). The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *Journal of Neuroscience Methods*, *264*, 47–56. https://doi.org/10.1016/j.jneumeth.2016.03.001

Liang, S., Yu, W., Ma, X., Luo, S., Zhang, J., Sun, X., Luo, X., & Zhang, Y. (2020). Psychometric properties of the MATRICS Consensus Cognitive Battery (MCCB) in Chinese patients with major depressive disorder. *Journal of Affective Disorders*, *265*, 132–138. https://doi.org/10.1016/j.jad.2020.01.052

Lim, J. H., Oh, I. K., Han, C., Huh, Y. J., Jung, I. K., Patkar, A. A., Steffens, D. C., & Jang, B. H. (2013). Sensitivity of cognitive tests in four cognitive domains in discriminating MDD patients from healthy controls: A meta-analysis. *International Psychogeriatrics*, *25*(9), 1543–1557. https://doi.org/10.1017/S1041610213000689

Lima, F., Rabelo-da-Ponte, F. D., Bücker, J., Czepielewski, L., Hasse-Sousa, M., Telesca, R., Sole, B., Reinares, M., Vieta, E., & Rosa, A. R. (2019). Identifying cognitive subgroups in bipolar disorder: A cluster analysis. *Journal of Affective Disorders*, *246*, 252–261. https://doi.org/10.1016/j.jad.2018.12.044

Lin, J., Su, Y., Shi, C., Liu, Q., Wang, G., Wei, J., Zhu, G., Chen, Q., Tian, H., Zhang, K., Wang, X., Zhang, N., Wang, Y., Yu, X., & Si, T. (2021). Neurocognitive profiles of patients with first-episode and recurrent depression: a cross-sectional comparative study from China. *Journal of Affective Disorders*, *286*, 110–116. https://doi.org/10.1016/j.jad.2021.02.068

Lin, X., Lu, D., Huang, Z., Chen, W., Luo, X., & Zhu, Y. (2019). The associations between subjective and objective cognitive functioning across manic or hypomanic, depressed, and euthymic states in Chinese bipolar patients. *Journal of Affective Disorders*, *249*, 73–81. https://doi.org/10.1016/j.jad.2019.02.025

Little, B. (2023). Future approaches to investigating cognitive functioning in Bipolar Disorder. *Bipolar Disorders*, *25*, 158–159. https://doi.org/10.1111/bdi.13268

Little, B., & Gallagher, P. (2020). Tests of processing speed and sustained attention in people with mood disorders: a systematic review protocol. In *PROSPERO* (p. CRD42020179443). PROPSPERO.

Liu, J., Dong, Q., Lu, X., Sun, J., Zhang, L., Wang, M., Wan, P., Guo, H., Zhao, F., Ju, Y., Yan, D., Li, H., Fang, H., Guo, W., Liao, M., Zhang, X., Zhang, Y., Liu, B., & Li, L. (2019). Exploration of Major Cognitive Deficits in Medication-Free Patients With Major Depressive Disorder. *Frontiers in Psychiatry*, *10*, 1–8. https://doi.org/10.3389/fpsyt.2019.00836

Liu, T., Zhong, S., Wang, B., Liao, X., Lai, S., & Jia, Y. (2019). Similar profiles of cognitive domain deficits between medication-naïve patients with bipolar II depression and those with major depressive disorder. *Journal of Affective Disorders*, *243*, 55–61. https://doi.org/10.1016/j.jad.2018.05.040

Lloyd, A. J., Moore, P. B., Cousins, D. A., Thompson, J. M., McAllister, V. L., Hughes, J. H., Ferrier, I. N.,

& Young, A. H. (2009). White matter lesions in euthymic patients with bipolar disorder. *Acta Psychiatrica Scandinavica*, *120*(6), 481–491. https://doi.org/10.1111/j.1600-0447.2009.01416.x

Lu, Z., Wang, Y., & Xun, G. (2021). Neurocognition Function of Patients With Bipolar Depression, Unipolar Depression, and Depression With Bipolarity. *Frontiers in Psychiatry*, *12*, 696903. https://doi.org/10.3389/fpsyt.2021.696903

Luperdi, S. C., Correa-Ghisays, P., Vila-Francés, J., Selva-Vera, G., Salazar-Fraile, J., Cardoner, N., Ruiz-Veguilla, M., Livianos, L., Tabarés-Seisdedos, R., & Balanzá-Martínez, V. (2021). Is processing speed a valid neurocognitive endophenotype in bipolar disorder? Evidence from a longitudinal, family study. *Journal of Psychiatric Research*, *141*, 241–247. https://doi.org/10.1016/j.jpsychires.2021.07.008

Lyness, S. A., Eaton, E. M., & Schneider, L. S. (1994). Cognitive performance in older and middle-aged depressed outpatients and controls. *Journals of Gerontology*, *49*(3), 129–136. https://doi.org/10.1093/geronj/49.3.P129

Maalouf, F. T., Klein, C., Clark, L., Sahakian, B. J., LaBarbara, E. J., Versace, A., Hassel, S., Almeida, J. R. C., & Phillips, M. L. (2010). Impaired sustained attention and executive dysfunction: Bipolar disorder versus depression-specific markers of affective disorders. *Neuropsychologia*, *48*(6), 1862–1868. https://doi.org/10.1016/j.neuropsychologia.2010.02.015

Machado, A., Barroso, J., Molina, Y., Nieto, A., Díaz-Flores, L., Westman, E., & Ferreira, D. (2018). Proposal for a hierarchical, multidimensional, and multivariate approach to investigate cognitive aging. *Neurobiology of Aging*, *71*, 179–188. https://doi.org/10.1016/j.neurobiolaging.2018.07.017

Macoveanu, J., Freeman, K. O., Kjærstad, H. L., Knudsen, G. M., Kessing, L. V., & Miskowiak, K. W. (2021). Structural brain abnormalities associated with cognitive impairments in bipolar disorder. *Acta Psychiatrica Scandinavica*, *144*(4), 379–391. https://doi.org/10.1111/acps.13349

Macoveanu, J., Mariegaard, J., Petersen, J. Z., Fisher, P. M., Vinberg, M., Jørgensen, M. B., Knudsen, G. M., Kessing, L. V., & Miskowiak, K. W. (2023). Neural basis of memory impairments and relation to functional disability in fully or partially remitted patients with affective disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 110767. https://doi.org/10.1016/J.PNPBP.2023.110767

Macoveanu, J., Petersen, J. Z., Fisher, P. M., Kessing, L. V., Knudsen, G. M., & Miskowiak, K. W. (2023). Associations between aberrant working memory-related neural activity and cognitive impairments in somatically healthy, remitted patients with mood disorders. *Psychological Medicine*, 1–11. https://doi.org/10.1017/S0033291723000715

MacQueen, G. M., & Memedovich, K. A. (2017). Cognitive dysfunction in major depression and bipolar disorder: Assessment and treatment options. *Psychiatry and Clinical Neurosciences*, *71*, 18–27. https://doi.org/10.1111/pcn.12463

Madan, C. R., & Kensinger, E. A. (2018). Predicting age from cortical structure across the lifespan. *European Journal of Neuroscience*, *47*(5), 399–416. https://doi.org/10.1111/ejn.13835

Mahlberg, R., Adli, M., Bschor, T., & Kienast, T. (2008). Age effects on trail making test during acute depressive and manic episode. *International Journal of Neuroscience*, *118*(9), 1347–1356. https://doi.org/10.1080/00207450601059452

Majer, M., Ising, M., Künzel, H., Binder, E. B., Holsboer, F., Modell, S., & Zihl, J. (2004). Impaired divided attention predicts delayed response and risk to relapse in subjects with depressive disorders. *Psychological Medicine*, *34*(8), 1453–1463. https://doi.org/10.1017/S0033291704002697

Malhi, G. S., Bassett, D., Boyce, P., Bryant, R., Fitzgerald, P. B., Fritz, K., Hopwood, M., Lyndon, B., Mulder, R., Murray, G., Porter, R., & Singh, A. B. (2015). Royal Australian and New Zealand

College of Psychiatrists clinical practice guidelines for mood disorders. *Australian and New Zealand Journal of Psychiatry*, *49*(12), 1087–1206. https://doi.org/10.1177/0004867415617657

Malhi, G. S., Ivanovski, B., Hadzi-Pavlovic, D., Mitchell, P. B., Vieta, E., & Sachdev, P. (2007). Neuropsychological deficits and functional impairment in bipolar depression, hypomania and euthymia. *Bipolar Disorders*, *9*(1–2), 114–125. https://doi.org/10.1111/j.1399-5618.2007.00324.x

Mandelli, L., Petrelli, C., & Serretti, A. (2015). The role of specific early trauma in adult depression: A meta-analysis of published literature. Childhood trauma and adult depression. *European Psychiatry*, *30*(6), 665–680. https://doi.org/10.1016/j.eurpsy.2015.04.007

Mann-Wrobel, M. C., Carreno, J. T., & Dickinson, D. (2011). Meta-analysis of neuropsychological functioning in euthymic bipolar disorder: An update and investigation of moderator variables. *Bipolar Disorders*, *13*(4), 334–342. https://doi.org/10.1111/j.1399-5618.2011.00935.x

Marotta, A., Chiaie, R. D., Spagna, A., Bernabei, L., Sciarretta, M., Roca, J., Biondi, M., & Casagrande, M. (2015). Impaired conflict resolution and vigilance in euthymic bipolar disorder. *Psychiatry Research*, *229*(1–2), 490–496. https://doi.org/10.1016/j.psychres.2015.06.026

Massalha, Y., Maggioni, E., Callari, A., Brambilla, P., & Delvecchio, G. (2023). A review of resting-state fMRI correlations with executive functions and social cognition in bipolar disorder. *Journal of Affective Disorders*, *334*, 337–351. https://doi.org/10.1016/j.jad.2023.03.084

Masuda, Y., Okada, G., Takamura, M., Shibasaki, C., Yoshino, A., Yokoyama, S., Ichikawa, N., Okuhata, S., Kobayashi, T., Yamawaki, S., & Okamoto, Y. (2020). White matter abnormalities and cognitive function in euthymic patients with bipolar disorder and major depressive disorder. *Brain and Behavior*, *10*(12), 1–11. https://doi.org/10.1002/brb3.1868

Mazoyer, B., Zago, L., Mellet, E., Bricogne, S., Etard, O., Houde, O., Crivello, F., Joliot, M., Petit, L., & Tzourio-Mazoyer, N. (2001). Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Research Bulletin*, *54*(3), 287–298.

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

McIntosh, A., Moorhead, T. W. J., McKirdy, J., Hall, J., Sussmann, J. E. D., Stanfield, A. C., Harris, J. M., Johnstone, E. C., & Lawrie, S. M. (2009). Prefrontal gyral folding and its cognitive correlates in bipolar disorder and schizophrenia. *Acta Psychiatrica Scandinavica*, *119*(3), 192–198. https://doi.org/10.1111/j.1600-0447.2008.01286.x

McIntosh, A. R., & Mišić, B. (2013). Multivariate statistical analyses for neuroimaging data. *Annual Review of Psychology*, *64*, 499–525. https://doi.org/10.1146/annurev-psych-113011-143804

McIntyre, R. S., Best, M. W., Bowie, C. R., Carmona, N. E., Cha, D. S., Lee, Y., Subramaniapillai, M., Mansur, R. B., Barry, H., Baune, B. T., Culpepper, L., Fossati, P., Greer, T. L., Harmer, C., Klag, E., Lam, R. W., H.-U, W., & Harrison, J. (2017). The THINC-integrated tool (THINC-it) screening assessment for cognitive dysfunction: Validation in patients with major depressive disorder. *Journal of Clinical Psychiatry*, *78*(7), 873–881.

McPhilemy, G., Nabulsi, L., Kilmartin, L., O'Hora, D., O'Donoghue, S., Tronchin, G., Costello, L., Najt, P., Ambati, S., Neilsen, G., Creighton, S., Byrne, F., McLoughlin, J., McDonald, C., Hallahan, B., & Cannon, D. M. (2020). Neuroanatomical Dysconnectivity Underlying Cognitive Deficits in Bipolar Disorder. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *5*(2), 152–162. https://doi.org/10.1016/j.bpsc.2019.09.004

McPhilemy, G., Nabulsi, L., Kilmartin, L., Whittaker, J. R., Martyn, F. M., Hallahan, B., McDonald, C., Murphy, K., & Cannon, D. M. (2020). Resting-State Network Patterns Underlying Cognitive Function in Bipolar Disorder: A Graph Theoretical Analysis. *Brain Connectivity*, *10*(7), 355–367. https://doi.org/10.1089/brain.2019.0709

Meda, S. A., Ruaño, G., Windemuth, A., O'Neil, K., Berwise, C., Dunn, S. M., Boccaccio, L. E., Narayanan, B., Kocherla, M., Sprooten, E., Keshavan, M. S., Tamminga, C. A., Sweeney, J. A., Clementz, B. A., Calhoun, V. D., & Pearlson, G. D. (2014). Multivariate analysis reveals genetic associations of the resting default mode network in psychotic bipolar disorder and schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(19), 2066–2075. https://doi.org/10.1073/pnas.1313093111

Menkes, M. W., Armstrong, K., Blackford, J. U., Heckers, S., & Woodward, N. D. (2019). Neuropsychological functioning in early and chronic stages of schizophrenia and psychotic bipolar disorder. *Schizophrenia Research*, *206*, 413–419. https://doi.org/10.1016/j.schres.2018.10.009

Merikangas, K. R., Jin, R., He, J. P., Kessler, R. C., Lee, S., Sampson, N. A., Viana, M. C., Andrade, L. H., Hu, C., Karam, E. G., Ladea, M., Medina-Mora, M. E., Ono, Y., Posada-Villa, J., Sagar, R., Wells, J. E., & Zarkov, Z. (2011). Prevalence and correlates of bipolar spectrum disorder in the World Mental Health Survey Initiative. *Archives of General Psychiatry*, *68*(3), 241–251. https://doi.org/10.1001/archgenpsychiatry.2011.12

Mesbah, R., Koenders, M. A., A van der Wee, N. J., Giltay, E. J., van Hemert, A. M., de Leeuw, M., der Wee, V., & Author, C. (2023). Association Between the Fronto-Limbic Network and Cognitive and Emotional Functioning in Individuals With Bipolar Disorder: A Systematic Review and Meta-analysis. *JAMA Psychiatry*, *80*(5), 432–440. https://doi.org/10.1001/JAMAPSYCHIATRY.2023.0131

Meyers, J. E. (2017). Trail Making Test. In J. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of Clinical Neuropsychology*. Springer, Cham. https://doi.org/10.1007/978-3-319-56782-2

Mihalik, A., Chapman, J., Adams, R. A., Winter, N. R., Ferreira, F. S., Shawe-Taylor, J., & Maurao-Miranda, J. (2022). Canonical Correlation Analysis and Partial Least Squares for identifying brain-behaviour associations: a tutorial and a comparative study. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *In Press*. https://doi.org/10.1016/j.bpsc.2022.07.012

Mikolajewicz, N., & Komarova, S. V. (2019). Meta-analytic methodology for basic research: A practical guide. *Frontiers in Physiology*, *10*. https://doi.org/10.3389/fphys.2019.00203

Milas, D. Ž., & Milas, G. (2019). Working memory in patients with schizophrenia and bipolar affective disorder: Quantitative or qualitative differences? *Psychiatria Danubina*, *31*(1), 54–61. https://doi.org/10.24869/psyd.2019.54

Misdraji, E. L., & Gass, C. S. (2010). The Trail Making Test and its neurobehavioral components. *Journal of Clinical and Experimental Neuropsychology*, *32*(2), 159–163. https://doi.org/10.1080/13803390902881942

Miskowiak, K. W., Burdick, K. E., Martinez-Aran, A., Bonnin, C. M., Bowie, C. R., Carvalho, A. F., Gallagher, P., Lafer, B., López-Jaramillo, C., Sumiyoshi, T., McIntyre, R. S., Schaffer, A., Porter, R. J., Purdon, S., Torres, I. J., Yatham, L. N., Young, A. H., Kessing, L. V., & Vieta, E. (2018). Assessing and addressing cognitive impairment in bipolar disorder: the International Society for Bipolar Disorders Targeting Cognition Task Force recommendations for clinicians. *Bipolar Disorders*, *20*(3), 184–194. https://doi.org/10.1111/bdi.12595

Miskowiak, K. W., Burdick, K. E., Martinez-Aran, A., Bonnin, C. M., Bowie, C. R., Carvalho, A. F., Gallagher, P., Lafer, B., López-Jaramillo, C., Sumiyoshi, T., McIntyre, R. S., Schaffer, A., Porter, R. J., Torres, I. J., Yatham, L. N., Young, A. H., Kessing, L. V., & Vieta, E. (2017). Methodological recommendations for cognition trials in bipolar disorder by the International Society for Bipolar Disorders Targeting Cognition Task Force. *Bipolar Disorders*, *19*(8), 614–626. https://doi.org/10.1111/bdi.12534

Miskowiak, K. W., Kjærstad, H. L., Lemvigh, C. K., Ambrosen, K. S., Thorvald, M. S., Kessing, L. V., Glenthoj, B. Y., Ebdrup, B. H., & Fagerlund, B. (2023). Neurocognitive subgroups among newly

diagnosed patients with schizophrenia spectrum or bipolar disorders: A hierarchical cluster analysis. *Journal of Psychiatric Research*. https://doi.org/10.1016/j.jpsychires.2023.05.025

Miskowiak, K. W., Kjærstad, H. L., Meluken, I., Petersen, J. Z., Maciel, B. R., Köhler, C. A., Vinberg, M., Kessing, L. V., & Carvalho, A. F. (2017). The search for neuroimaging and cognitive endophenotypes: A critical systematic review of studies involving unaffected first-degree relatives of individuals with bipolar disorder. *Neuroscience and Biobehavioral Reviews*, *73*, 1–22. https://doi.org/10.1016/j.neubiorev.2016.12.011

Miskowiak, K. W., & Petersen, C. S. (2019). Neuronal underpinnings of cognitive impairment and -improvement in mood disorders. *CNS Spectrums*, *24*(1), 30–53. https://doi.org/10.1017/S1092852918001062

Miskowiak, K. W., Seeberg, I., Jensen, M., Balanzá-Martínez, V., del Mar Bonnin, C., Bowie, C. R., Carvalho, A. F., Dols, A., Douglas, K. M., Gallagher, P., Hasler, G., Lafer, B., Lewandowski, K. E., López-Jaramillo, C., Martinez-Aran, A., McIntyre, R. S., Porter, R. J., Purdon, S. E., Schaffer, A., … Vieta, E. (2022). Randomised controlled cognition trials in remitted patients with mood disorders published between 2015 and 2021: A systematic review by the International Society for Bipolar Disorders Targeting Cognition Task Force. *Bipolar Disorders*, *24*, 354–374. https://doi.org/10.1111/bdi.13193

Miskowiak, K. W., Vinberg, M., Christensen, E. M., & Kessing, L. V. (2012). Is there a difference in subjective experience of cognitive function in patients with unipolar disorder versus bipolar disorder? *Nordic Journal of Psychiatry*, *66*(6), 389–395. https://doi.org/10.3109/08039488.2012.658862

Miskowiak, K. W., Vinberg, M., Glerup, L., Paulson, O. B., Knudsen, G. M., Ehrenreich, H., Harmer, C. J., Kessing, L. V., Siebner, H. R., & Macoveanu, J. (2016). Neural correlates of improved executive function following erythropoietin treatment in mood disorders. *Psychological Medicine*, *46*(8), 1679–1691. https://doi.org/10.1017/S0033291716000209

Miskowiak, K. W., Yalin, N., Seeberg, I., Burdick, K. E., Balanzá-Martínez, V., Bonnin, C. del M., Bowie, C. R., Carvalho, A. F., Dols, A., Douglas, K. M., Gallagher, P., Hasler, G., Kessing, L. V., Lafer, B., Lewandowski, K. E., López-Jaramillo, C., Martinez-Aran, A., McIntyre, R. S., Porter, R. J., … Stokes, P. R. A. (2022). Can magnetic resonance imaging enhance the assessment of potential new treatments for cognitive impairment in mood disorders? A systematic review and position paper by the International Society for Bipolar Disorders Targeting Cognition Task Force. *Bipolar Disorders*, *In Press*. https://doi.org/10.1111/bdi.13247

Miyake, A., Emerson, M. J., & Friedman, N. P. (2000). Assessment of Executive Functions in Clinical Settings: Problems and Recommendations. *Seminars in Speech and Language*, *21*(2), 169–183.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Miyata, A., Iwamoto, K., Kawano, N., Aleksic, B., Ando, M., Ebe, K., Fujita, K., Yokoyama, M., Akiyama, T., Igarashi, Y., & Ozaki, N. (2018). Driving performance of stable outpatients with depression undergoing real-world treatment. *Psychiatry and Clinical Neurosciences*, *72*(6), 399–408. https://doi.org/10.1111/pcn.12648

Modabbernia, A., Janiri, D., Doucet, G. E., Reichenberg, A., & Frangou, S. (2021). Multivariate Patterns of Brain-Behavior-Environment Associations in the Adolescent Brain and Cognitive Development Study. *Biological Psychiatry*, *89*(5), 510–520. https://doi.org/10.1016/j.biopsych.2020.08.014

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., Estarli, M., Barrera, E. S. A., Martínez-Rodríguez, R., Baladia, E., Agüero, S. D., Camacho, S.,

Buhring, K., Herrero-López, A., Gil-González, D. M., Altman, D. G., Booth, A., … Whitlock, E. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, *4*, 1. https://doi.org/10.1186/2046-4053-4-1

Moss, R. A., Finkelmeyer, A., Robinson, L. J., Thompson, J. M., Watson, S., Ferrier, I. N., & Gallagher, P. (2016). The impact of target frequency on intra-individual variability in euthymic bipolar disorder: A comparison of two sustained attention tasks. *Frontiers in Psychiatry*, *7*(August), 106. https://doi.org/10.3389/fpsyt.2016.00106

Naim-Feil, J., Bradshaw, J. L., Sheppard, D. M., Rosenberg, O., Levkovitz, Y., Dannon, P., Fitzgerald, P. B., Isserles, M., & Zangen, A. (2016). Neuromodulation of attentional control in major depression: A pilot DeepTMS study. *Neural Plasticity*, *2016*. https://doi.org/10.1155/2016/5760141

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Nebes, R. D., Butters, M. A., Mulsant, B. H., Pollock, B. G., Zmuda, M. D., Houck, P. R., & Reynolds, C. F. (2000). Decreased working memory and processing speed mediate cognitive impairment in geriatric depression. *Psychological Medicine*, *30*(3), 679–691. https://doi.org/10.1017/S0033291799001968

Necus, J., Sinha, N., Smith, F. E., Thelwall, P. E., Flowers, C. J., Taylor, P. N., Blamire, A. M., Cousins, D. A., & Wang, Y. (2019). White matter microstructural properties in bipolar disorder in relationship to the spatial distribution of lithium in the brain. *Journal of Affective Disorders*, *253*, 224–231. https://doi.org/10.1016/j.jad.2019.04.075

Necus, J., Smith, F. E., Thelwall, P. E., Flowers, C. J., Sinha, N., Taylor, P. N., Blamire, A. M., Wang, Y., & Cousins, D. A. (2021). Quantification of brain proton longitudinal relaxation (T1) in lithium-treated and lithium-naïve patients with bipolar disorder in comparison to healthy controls. *Bipolar Disorders*, *23*, 41–48. https://doi.org/10.1111/bdi.12878

Nehra, R., Chakrabarti, S., Pradhan, B. K., & Khehra, N. (2006). Comparison of cognitive functions between first- and multi-episode bipolar affective disorders. *Journal of Affective Disorders*, *93*(1–3), 185–192. https://doi.org/10.1016/j.jad.2006.03.013

Nehra, R., Grover, S., Sharma, S., Sharma, A., & Sarkar, S. (2014). Neuro-cognitive functioning in unaffected siblings of patients with bipolar disorder: Comparison with bipolar patients and healthy controls. *Indian Journal of Psychiatry*, *56*(3), 283–288. https://doi.org/10.4103/0019-5545.140645

Nelson, H. (1982). *National Adult Reading Test (NART)*. Nelson Publishing Company.

Nguyen, T. T., Kovacevic, S., Dev, S. I., Lu, K., Liu, T. T., & Eyler, L. T. (2017). Dynamic functional connectivity in bipolar disorder is associated with executive function and processing speed: A preliminary study. *Neuropsychology*, *31*(1), 73–83. https://doi.org/10.1037/neu0000317

Nilsson, J., Thomas, A. J., O'Brien, J. T., & Gallagher, P. (2014). White matter and cognitive decline in aging: A focus on processing speed and variability. *Journal of the International Neuropsychological Society*, *20*(3), 262–267. https://doi.org/10.1017/S1355617713001458

Nilsson, J., Thomas, A. J., Stevens, L. H., McAllister-Williams, R. H., Ferrier, I. N., & Gallagher, P. (2016). The interrelationship between attentional and executive deficits in major depressive disorder. *Acta Psychiatrica Scandinavica*, *134*(1), 73–82. https://doi.org/10.1111/acps.12570

Normala, I., Abdul Hamid, A. R., Azlin, B., Nik Ruzyanei, N. J., Hazli, Z., & Shah, S. A. (2010). Executive function and attention span in euthymic patients with bipolar 1 disorder. *Medical Journal of Malaysia*, *65*(3), 192–196. https://doi.org/10.1016/s1876-2018(11)60194-9

Nuechterlein, K. H., Green, M. F., Kern, R. S., Baade, L. E., Barch, D. M., Cohen, J. D., Essock, S., Fenton, W. S., Frese, F. J., Gold, J. M., Goldberg, T., Heaton, R. K., Keefe, R. S. E., Kraemer, H., Mesholam-Gately, R., Seidman, L. J., Stover, E., Weinberger, D. R., Young, A. S., … Marder, S. R. (2008). The MATRICS consensus cognitive battery, part 1: Test selection, reliability, and validity. *American Journal of Psychiatry*, *165*(2), 203–213. https://doi.org/10.1176/appi.ajp.2007.07010042

Nunes, A., Schnack, H. G., Ching, C. R. K., Agartz, I., Akudjedu, T. N., Alda, M., Alnæs, D., Alonso-Lana, S., Bauer, J., Baune, B. T., Bøen, E., Bonnin, C. del M., Busatto, G. F., Canales-Rodríguez, E. J., Cannon, D. M., Caseras, X., Chaim-Avancini, T. M., Dannlowski, U., Díaz-Zuluaga, A. M., … Hajek, T. (2020). Using structural MRI to identify bipolar disorders – 13 site machine learning study in 3020 individuals from the ENIGMA Bipolar Disorders Working Group. *Molecular Psychiatry*, *25*(9), 2130–2143. https://doi.org/10.1038/s41380-018-0228-9

Oakley, B., Loth, E., & Murphy, D. G. (2021). Autism and mood disorders. *International Review of Psychiatry*, *33*(3), 280–299. https://doi.org/10.1080/09540261.2021.1872506

Oertel-Knöchel, V., Reuter, J., Reinke, B., Marbach, K., Feddern, R., Alves, G., Prvulovic, D., Linden, D. E. J., & Knöchel, C. (2015). Association between age of disease-onset, cognitive performance and cortical thickness in bipolar disorders. *Journal of Affective Disorders*, *174*, 627–635. https://doi.org/10.1016/j.jad.2014.10.060

Okasha, T. A., El Sheikh, M. M., El Missiry, A. A., El Missiry, M. A., El Serafi, D., El Kholy, S., & Abdel Aziz, K. (2014). Cognitive functions in euthymic Egyptian patients with bipolar disorder: Are they different from healthy controls? *Journal of Affective Disorders*, *166*, 14–21. https://doi.org/10.1016/j.jad.2014.04.051

Oliveira, J. P. de, Jansen, K., Cardoso, T. de A., Mondin, T. C., Souza, L. D. de M., Silva, R. A. da, & Pedrotti Moreira, F. (2021). Predictors of conversion from major depressive disorder to bipolar disorder. *Psychiatry Research*, *297*, 113740. https://doi.org/10.1016/j.psychres.2021.113740

Ortiz, A. C. D., Upeguí, C. V., Ospina, J. P. Z., Acevedo, D. C. A., Zapata, J. A. P., & LópezJaramillo, C. A. (2021). Correlation Between Cognitive Performance and Structural Neuroanatomy in Patients with Type I Bipolar Affective Disorder Treated with and Without Lithium. *Revista Colombiana de Psiquiatría*.

Osborne, J. W. (2013). *Best Practices in Data Cleaning: A Complete Guide to Everything You Need to Do Before and After Collecting Your Data Title*. Sage Publications, Inc.

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan — a web and mobile app for systematic reviews. *Systematic Reviews*, *5*, 210. https://doi.org/10.1186/s13643-016-0384-4

Paelecke-Habermann, Y., Pohl, J., & Leplow, B. (2005). Attention and executive functions in remitted major depression patients. *Journal of Affective Disorders*, *89*(1–3), 125–135. https://doi.org/10.1016/j.jad.2005.09.006

Parkinson, W. L., Rehman, Y., Rathbone, M., & Upadhye, S. (2020). Performances on individual neurocognitive tests by people experiencing a current major depression episode: A systematic review and meta-analysis. *Journal of Affective Disorders*, *276*, 249–259. https://doi.org/10.1016/j.jad.2020.07.036

Pattanayak, R. D., Sagar, R., & Mehta, M. (2012). Neuropsychological performance in euthymic Indian patients with bipolar disorder type I: Correlation between quality of life and global functioning. *Psychiatry and Clinical Neurosciences*, *66*(7), 553–563. https://doi.org/10.1111/j.1440-1819.2012.02400.x

Pearson, O., Uglik-Marucha, N., Miskowiak, K. W., Cairney, S. A., Rosenzweig, I., Young, A. H., & Stokes, P. R. A. (2021). The relationship between sleep disturbance and cognitive impairment in

mood disorders: A systematic review. *MedRxiv*, 2021–08. https://doi.org/10.1101/2021.08.12.21261470

Penttilä, J., Paillère-Martino, M. L., Martinot, J. L., Ringuenet, D., Wessa, M., Houenou, J., Gallarda, T., Bellivier, F., Galinowski, A., Bruguière, P., Pinabel, F., Leboyer, M., Olié, J. P., Duchesnay, E., Artiges, E., Mangin, J. F., & Cachia, A. (2009). Cortical folding in patients with bipolar disorder or unipolar depression. *Journal of Psychiatry and Neuroscience*, *34*(2), 127–135.

Piani, M. C., Maggioni, E., Delvecchio, G., & Brambilla, P. (2022). Sustained attention alterations in major depressive disorder: A review of fMRI studies employing Go/No-Go and CPT tasks. *Journal of Affective Disorders*, *303*, 98–113. https://doi.org/10.1016/J.JAD.2022.02.003

Pier, M. P. B. I., Hulstijn, W., & Sabbe, B. G. C. (2004). Differential patterns of psychomotor functioning in unmedicated melancholic and nonmelancholic depressed patients. *Journal of Psychiatric Research*, *38*(4), 425–435. https://doi.org/10.1016/j.jpsychires.2003.11.008

Pillai, K. C. S. (1955). Some New Test Criteria in Multivariate Analysis. *The Annals of Mathematical Statistics*, *26*(1), 117–121. https://doi.org/10.1214/aoms/1177728599

Pol, H. E. H., Schnack, H. G., Posthuma, D., Mandl, R. C. W., Baaré, W. F., Van Oel, C., Van Haren, N. E., Collins, D. L., Evans, A. C., Amunts, K., Bürgel, U., Zilles, K., De Geus, E., Boomsma, D. I., & Kahn, R. S. (2006). Genetic contributions to human brain morphology and intelligence. *Journal of Neuroscience*, *26*(40), 10235–10242. https://doi.org/10.1523/JNEUROSCI.1312-06.2006

Poletti, S., Aggio, V., Brioschi, S., Dallaspezia, S., Colombo, C., & Benedetti, F. (2017). Multidimensional cognitive impairment in unipolar and bipolar depression and the moderator effect of adverse childhood experiences. *Psychiatry and Clinical Neurosciences*, *71*(5), 309–317. https://doi.org/10.1111/pcn.12497

Poletti, S., Bollettini, I., Mazza, E., Locatelli, C., Radaelli, D., Vai, B., Smeraldi, E., Colombo, C., & Benedetti, F. (2015). Cognitive performances associate with measures of white matter integrity in bipolar disorder. *Journal of Affective Disorders*, *174*, 342–352. https://doi.org/10.1016/j.jad.2014.12.030

Poletti, S., Papa, G. S., Locatelli, C., Colombo, C., & Benedetti, F. (2014). Neuropsychological deficits in bipolar depression persist after successful antidepressant treatment. *Journal of Affective Disorders*, *156*, 144–149. https://doi.org/10.1016/j.jad.2013.11.023

Porter, R. J., Gallagher, P., Thompson, J. M., & Young, A. H. (2003). Neurocognitive impairment in drug-free patients with major depressive disorder. *British Journal of Psychiatry*, *182*(MAR.), 214–220. https://doi.org/10.1192/bjp.182.3.214

Porter, R. J., Robinson, L. J., Malhi, G. S., & Gallagher, P. (2015). The neurocognitive profile of mood disorders - a review of the evidence and methodological issues. *Bipolar Disorders*, *17*, 21–40. https://doi.org/10.1111/bdi.12342

Pradhan, B. K., Chakrabarti, S., Nehra, R., & Mankotia, A. (2008). Cognitive functions in bipolar affective disorder and schizophrenia: Comparison. *Psychiatry and Clinical Neurosciences*, *62*(5), 515–525. https://doi.org/10.1111/j.1440-1819.2008.01844.x

Proctor, R. W., & Schneider, D. W. (2017). Hick's Law for Choice Reaction Time: A Review. *The Quarterly Journal of Experimental Psychology*, *71*(6), 1281–1299. https://doi.org/10.1080/17470218.2017.1322622

Psychological Corporation. (1998). *Vigil™ Continuous Performance Test*. Harcourt Brace & Company.

Purdon, S. (2005). *The Screen for Cognitive Impairment in Psychiatry: Administration and Psychometric Properties*. PNL.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Radwan, D. N., Okasha, T., Elmissiry, M., Sadek, H., Khalifa, A., & Abdelaziz, K. (2013). Cognitive impairment in Egyptian euthymic patients with bipolar i disorder compared with controls. *Middle East Current Psychiatry*, *20*(4), 197–204. https://doi.org/10.1097/01.XME.0000433325.69290.c9

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 676–682. https://doi.org/10.1073/pnas.98.2.676

Rashid, B., Damaraju, E., Pearlson, G. D., & Calhoun, V. D. (2014). Dynamic connectivity states estimated from resting fMRI Identify differences among schizophrenia, bipolar disorder, and healthy control subjects. *Frontiers in Human Neuroscience*, *8*, 897. https://doi.org/10.3389/fnhum.2014.00897

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, *86*(3), 446–461. https://doi.org/10.1037/0033-2909.86.3.446

Rathbone, J., Carter, M., Hoffmann, T., & Glasziou, P. (2015). Better duplicate detection for systematic reviewers: Evaluation of Systematic Review Assistant-Deduplication Module. *Systematic Reviews*, *4*, 1–6. https://doi.org/10.1186/2046-4053-4-6

Raymond, M. R. (1986). Missing Data in Evaluation Research. *Evaluation & the Health Professions*, *9*(4), 395–420. https://doi.org/10.1177/016327878600900401

Raymond, M. R., & Roberts, D. M. (1987). A Comparison of Methods for Treating Incomplete Data in Selection Research. *Educational and Psychological Measurement*, *47*(1), 13–26. https://doi.org/10.1177/0013164487471002

Reitan, R. M. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual and Motor Skills*, *8*, 271–276.

Renner, F., Siep, N., Arntz, A., van de Ven, V., Peeters, F. P. M. L., Quaedflieg, C. W. E. M., & Huibers, M. J. H. (2017). Negative mood-induction modulates default mode network resting-state functional connectivity in chronic depression. *Journal of Affective Disorders*, *208*, 590–596. https://doi.org/10.1016/j.jad.2016.10.022

Rey, A. (1964). *L'Examen Clinique en Psychologie* (Press Universitaire de France (ed.)).

Rizk, M. M., Rubin-Falcone, H., Keilp, J., Miller, J. M., Sublette, M. E., Burke, A., Oquendo, M. A., Kamal, A. M., Abdelhameed, M. A., & Mann, J. J. (2017). White matter correlates of impaired attention control in major depressive disorder and healthy volunteers. *Journal of Affective Disorders*, *222*, 103–111. https://doi.org/10.1016/j.jad.2017.06.066

Robinson, L. J., Thompson, J. M., Gallagher, P., Goswami, U., Young, A. H., Ferrier, I. N., & Moore, P. B. (2006). A meta-analysis of cognitive deficits in euthymic patients with bipolar disorder. *Journal of Affective Disorders*, *93*, 105–115. https://doi.org/10.1016/j.jad.2006.02.016

Robinson, L. J., Thompson, J. M., Gallagher, P., Gray, J. M., Young, A. H., & Ferrier, I. N. (2013). Performance monitoring and executive control of attention in euthymic bipolar disorder: Employing the CPT-AX paradigm. *Psychiatry Research*, *210*(2), 457–464. https://doi.org/10.1016/j.psychres.2013.06.039

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

Rodrigue, A. L., McDowell, J. E., Tandon, N., Keshavan, M. S., Tamminga, C. A., Pearlson, G. D., Sweeney, J. A., Gibbons, R. D., & Clementz, B. A. (2018). Multivariate Relationships Between Cognition and Brain Anatomy Across the Psychosis Spectrum. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(12), 992–1002. https://doi.org/10.1016/j.bpsc.2018.03.012

Rodríguez-Cruces, R., Bernhardt, B. C., & Concha, L. (2020). Multidimensional associations between cognition and connectome organization in temporal lobe epilepsy. *NeuroImage*, *213*, 116706. https://doi.org/10.1016/j.neuroimage.2020.116706

Romero, E., Holtzman, J. N., Tannenhaus, L., Monchablon, R., Rago, C. M., Lolich, M., & Vázquez, G. H. (2016). Neuropsychological performance and affective temperaments in euthymic patients with bipolar disorder type II. *Psychiatry Research*, *238*, 172–180. https://doi.org/10.1016/j.psychres.2016.02.032

Ronold, E. H., Schmid, M. T., Oedegaard, K. J., & Hammar, Å. (2020). A Longitudinal 5-Year Follow-Up Study of Cognitive Function After First Episode Major Depressive Disorder: Exploring State, Scar and Trait Effects. *Frontiers in Psychiatry*, *11*(575867), 1–12. https://doi.org/10.3389/fpsyt.2020.575867

Rosenbaum, D., Haipt, A., Fuhr, K., Haeussinger, F. B., Metzger, F. G., Nuerk, H. C., Fallgatter, A. J., Batra, A., & Ehlis, A. C. (2017). Aberrant functional connectivity in depression as an index of state and trait rumination. *Scientific Reports*, *7*, 2174. https://doi.org/10.1038/s41598-017-02277-z

Rosenblat, J. D., Kakar, R., & McIntyre, R. S. (2016). The Cognitive Effects of Antidepressants in Major Depressive Disorder: A Systematic Review and Meta-Analysis of Randomized Clinical Trials. *The International Journal of Neuropsychopharmacology*, *19*(2). https://doi.org/10.1093/ijnp/pyv082

Roth, P. L. (1994). Missing Data: a Conceptual Review for Applied Psychologists. *Personnel Psychology*, *47*(3), 537–560. https://doi.org/10.1111/j.1744-6570.1994.tb01736.x

Rothman, K. J. (1990). No Adjustments Are Needed for Multiple Comparisons. *Epidemiology*, *1*(1), 43–46. https://doi.org/10.1097/00001648-199001000-00010

Roux, P., Etain, B., Cannavo, A.-S., Aubin, V., Aouizerate, B., Azorin, J.-M., Bellivier, F., Belzeaux, R., Bougerol, T., Cussac, I., Courtet, P., Kahn, J.-P., Leboyer, M., M'Bailara, K., Payet, M. P., Olié, E., Henry, C., & Passerieux, C. (2019). Prevalence and determinants of cognitive impairment in the euthymic phase of bipolar disorders: results from the FACE-BD cohort. *Psychological Medicine*, *49*, 519–527. https://doi.org/10.1017/s0033291718001186

Russo, M., Van Rheenen, T. E., Shanahan, M., Mahon, K., Perez-Rodriguez, M. M., Cuesta-Diaz, A., Larsen, E., Malhotra, A. K., & Burdick, K. E. (2017). Neurocognitive subtypes in patients with bipolar disorder and their unaffected siblings. *Psychological Medicine*, *47*(16), 2892–2905. https://doi.org/10.1017/S003329171700143X

Sachdev, P. S., Blacker, D., Blazer, D. G., Ganguli, M., Jeste, D. V., Paulsen, J. S., & Petersen, R. C. (2014). Classifying neurocognitive disorders: The DSM-5 approach. *Nature Reviews Neurology*, *10*(11), 634–642. https://doi.org/10.1038/nrneurol.2014.181

Salazar-Fraile, J., Balanzá-Martínez, V., Selva-Vera, G., Martínez-Arán, A., Sánchez-Moreno, J., Rubio, C., Vieta, E., Gómez-Beneyto, M., & Tabarés-Seisdedos, R. (2009). Motor speed predicts stability of cognitive deficits in both schizophrenic and bipolar I patients at one-year follow-up. *European Journal of Psychiatry*, *23*(3), 184–197. https://doi.org/10.4321/S0213-61632009000300007

Salthouse, T. A. (1996). The Processing-Speed Theory of Adult Age Differences in Cognition. *Psychological Review*, *103*(3), 403–428. https://doi.org/10.1037/0033-295X.103.3.403

Samamé, C., Szmulewicz, A. G., Valerio, M. P., Martino, D. J., & Strejilevich, S. A. (2017). Are major depression and bipolar disorder neuropsychologically distinct? A meta-analysis of comparative studies. *European Psychiatry*, *39*, 17–26. https://doi.org/10.1016/j.eurpsy.2016.06.002

Sánchez-Carro, Y., Portella, M. J., Leal-Leturia, I., Pujol, N. S., Etxandi, M., de Arriba-Arnau, A., Urretavizcaya, M., Pousa, E., Toll, A., Álvarez, P., Soria, V., & López-García, P. (2021). Age at illness onset and physical activity are associated with cognitive impairment in patients with

current diagnosis of major depressive disorder. *Journal of Affective Disorders*, *279*(October 2020), 343–352. https://doi.org/10.1016/j.jad.2020.10.032

Sánchez-Morla, E. M., López-Villarreal, A., Jiménez-López, E., Aparicio, A. I., Martínez-Vizcaíno, V., Roberto, R. J., Vieta, E., & Santos, J. L. (2019). Impact of number of episodes on neurocognitive trajectory in bipolar disorder patients: A 5-year follow-up study. *Psychological Medicine*, *49*(8), 1299–1307. https://doi.org/10.1017/S0033291718001885

Sarapas, C., Shankman, S. A., Harrow, M., & Faull, R. N. (2013). Attention/processing speed prospectively predicts social impairment 18 years later in mood disorders. *Journal of Nervous and Mental Disease*, *201*(9), 824–827. https://doi.org/10.1097/NMD.0b013e31829db5ed

Sax, K. W., Strakowski, S. M., Zimmerman, M. E., DelBello, M. P., Keck, P. E., & Hawkins, J. M. (1999). Frontosubcortical neuroanatomy and the continuous performance test in mania. *American Journal of Psychiatry*, *156*(1), 139–141. https://doi.org/10.1176/ajp.156.1.139

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177. https://doi.org/10.1037/1082-989X.7.2.147

Schmaal, L., Hibar, D. P., Sämann, P. G., Hall, G. B., Baune, B. T., Jahanshad, N., Cheung, J. W., Van Erp, T. G. M., Bos, D., Ikram, M. A., Vernooij, M. W., Niessen, W. J., Tiemeier, H., Hofman, A., Wittfeld, K., Grabe, H. J., Janowitz, D., Bülow, R., Selonke, M., … Veltman, D. J. (2017). Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. *Molecular Psychiatry*, *22*(6), 900–909. https://doi.org/10.1038/mp.2016.60

Schmaal, L., Veltman, D. J., Van Erp, T. G. M., Smann, P. G., Frodl, T., Jahanshad, N., Loehrer, E., Tiemeier, H., Hofman, A., Niessen, W. J., Vernooij, M. W., Ikram, M. A., Wittfeld, K., Grabe, H. J., Block, A., Hegenscheid, K., Völzke, H., Hoehn, D., Czisch, M., … Hibar, D. P. (2016). Subcortical brain alterations in major depressive disorder: Findings from the ENIGMA Major Depressive Disorder working group. *Molecular Psychiatry*, *21*(6), 806–812. https://doi.org/10.1038/mp.2015.69

Schmidt, G. J., Barbosa, A. O., de Assis, S. G., Nicaretta, D. H., & Schmidt, S. L. (2021). Attentional subdomains' deficits in Brazilian patients with major depressive episodes. *Neuropsychology*, *35*(2), 232–240. https://doi.org/10.1037/neu0000719

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H. M., & Wittmann, W. W. (2007). Individual Differences in Components of Reaction Time Distributions and Their Relations to Working Memory and Intelligence. *Journal of Experimental Psychology: General*, *136*(3), 414–429. https://doi.org/10.1037/0096-3445.136.3.414

Schubert, A. L., Hagemann, D., & Frischkorn, G. T. (2017). Is general intelligence little more than the speed of higher-order processing? *Journal of Experimental Psychology: General*, *146*(10), 1498–1512. https://doi.org/10.1037/xge0000325

Schumacher, J., Cromarty, R., Gallagher, P., Firbank, M. J., Thomas, A. J., Kaiser, M., Blamire, A. M., O'Brien, J. T., Peraza, L. R., & Taylor, J. P. (2019). Structural correlates of attention dysfunction in Lewy body dementia and Alzheimer's disease: an ex-Gaussian analysis. *Journal of Neurology*, *266*(7), 1716–1726. https://doi.org/10.1007/s00415-019-09323-y

Seli, P., Beaty, R. E., Marty-Dugas, J., & Smilek, D. (2019). Depression, anxiety, and stress and the distinction between intentional and unintentional mind wandering. *Psychology of Consciousness: Theory Research, and Practice*, *6*(2), 163–170. https://doi.org/10.1037/cns0000182

Semkovska, M., Quinlivan, L., O'Grady, T., Johnson, R., Collins, A., O'Connor, J., Knittle, H., Ahern, E., & Gload, T. (2019). Cognitive function following a major depressive episode: a systematic review and meta-analysis. *The Lancet Psychiatry*, *6*(10), 851–861.

https://doi.org/10.1016/S2215-0366(19)30291-3

Şentürk Cankorur, V., Demirel, H., & Atbaşoğlu, C. (2017). Cognitive functioning in euthymic bipolar patients on monotherapy with novel antipsychotics or mood stabilizers. *Noropsikiyatri Arsivi*, *54*(3), 244–250. https://doi.org/10.5152/npa.2017.15883

Sexton, C. E., McDermott, L., Kalu, U. G., Herrmann, L. L., Bradley, K. M., Allan, C. L., Le Masurier, M., MacKay, C. E., & Ebmeier, K. P. (2012). Exploring the pattern and neural correlates of neuropsychological impairment in late-life depression. *Psychological Medicine*, *42*(6), 1195–1202. https://doi.org/10.1017/S0033291711002352

Shan, C., Lee, S. Y., Chang, Y. H., Wu, J. Y. W., Chen, S. L., Chen, S. H., Hsiao, Y. L., Yang, H. F., Lee, I. H., Chen, P. S., Yeh, T. L., Yang, Y. K., & Lu, R. B. (2011). Neuropsychological functions in Han Chinese patients in Taiwan with bipolar II disorder comorbid and not comorbid with alcohol abuse/alcohol dependence disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *35*(1), 131–136. https://doi.org/10.1016/j.pnpbp.2010.10.004

Sheline, Y. I., Barch, D. M., Garcia, K., Gersing, K., Pieper, C., Welsh-Bohmer, K., Steffens, D. C., & Doraiswamy, P. M. (2006). Cognitive Function in Late Life Depression: Relationships to Depression Severity, Cerebrovascular Risk Factors and Processing Speed. *Biological Psychiatry*, *60*(1), 58–65. https://doi.org/10.1016/j.biopsych.2005.09.019

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences*, *44*(3), 535–551. https://doi.org/10.1016/j.paid.2007.09.015

Sherry, A., & Henson, R. K. (2005). Conducting and interpreting canonical correlation analysis in personality research: A user-friendly primer. *Journal of Personality Assessment*, *84*(1), 37–48. https://doi.org/10.1207/s15327752jpa8401_09

Shimizu, Y., Kitagawa, N., Mitsui, N., Fujii, Y., Toyomaki, A., Hashimoto, N., Kako, Y., Tanaka, T., Asakura, S., & Kusumi, I. (2013). Neurocognitive impairments and quality of life in unemployed patients with remitted major depressive disorder. *Psychiatry Research*, *210*(3), 913–918. https://doi.org/10.1016/j.psychres.2013.08.030

Smith, F. E., Thelwall, P. E., Necus, J., Flowers, C. J., Blamire, A. M., & Cousins, D. A. (2018). 3D 7Li magnetic resonance imaging of brain lithium distribution in bipolar disorder. *Molecular Psychiatry*, *23*(11), 2184–2191. https://doi.org/10.1038/s41380-018-0016-6

Smith, P. J., Need, A. C., Cirulli, E. T., Chiba-Falek, O., & Attix, D. K. (2013). A comparison of the Cambridge Automated Neuropsychological Test Battery (CANTAB) with "traditional" neuropsychological testing instruments. *Journal of Clinical and Experimental Neuropsychology*, *35*(3), 319–328. https://doi.org/10.1080/13803395.2013.771618

Smith, S. M., Nichols, T. E., Vidaurre, D., Winkler, A. M., Behrens, T. E. J., Glasser, M. F., Ugurbil, K., Barch, D. M., Van Essen, D. C., & Miller, K. L. (2015). A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nature Neuroscience*, *18*(11), 1565–1567. https://doi.org/10.1038/nn.4125

Smucny, J., Barch, D. M., Gold, J. M., Strauss, M. E., MacDonald, A. W., Boudewyn, M. A., Ragland, J. D., Silverstein, S. M., & Carter, C. S. (2019). Cross-diagnostic analysis of cognitive control in mental illness: Insights from the CNTRACS consortium. *Schizophrenia Research*, *208*, 377–383. https://doi.org/10.1016/j.schres.2019.01.018

Smucny, J., Lesh, T. A., Iosif, A. M., Niendam, T. A., Tully, L. M., & Carter, C. S. (2018). Longitudinal stability of cognitive control in early psychosis: Nondegenerative deficits across diagnoses. *Journal of Abnormal Psychology*, *127*(8), 781–788. https://doi.org/10.1037/abn0000356

Smucny, J., Lesh, T. A., Newton, K., Niendam, T. A., Ragland, J. D., & Carter, C. S. (2018). Levels of Cognitive Control: A Functional Magnetic Resonance Imaging-Based Test of an RDoC Domain

Across Bipolar Disorder and Schizophrenia. *Neuropsychopharmacology*, *43*(3), 598–606. https://doi.org/10.1038/npp.2017.233

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

Sobel, M. E. (1982). Asymptotic intervals for indirect effects in structural equations models. In S. Leinhart (Ed.), *Sociological methodology* (p. pp.290-312). Jossey-Bass.

Solé, B., Bonnin, C. M., Jiménez, E., Torrent, C., Torres, I., Varo, C., Valls, E., Montejo, L., Gómez-Ocaña, C., Tomioka, Y., Vieta, E., Martinez-Aran, A., & Reinares, M. (2018). Heterogeneity of functional outcomes in patients with bipolar disorder: a cluster-analytic approach. *Acta Psychiatrica Scandinavica*, *137*(6), 516–527. https://doi.org/10.1111/acps.12871

Solé, B., Bonnin, C. M., Torrent, C., Balanzá-Martínez, V., Tabarés-Seisdedos, R., Popovic, D., Martínez-Arán, A., & Vieta, E. (2012). Neurocognitive impairment and psychosocial functioning in bipolar II disorder. *Acta Psychiatrica Scandinavica*, *125*(4), 309–317. https://doi.org/10.1111/j.1600-0447.2011.01759.x

Soni, A., Singh, P., Shah, R., & Bagotia, S. (2017). Impact of cognition and clinical factors on functional outcome in patients with bipolar disorder. *East Asian Archives of Psychiatry*, *27*(1), 26–34.

Sørensen, H. J., Sæbye, D., Urfer-Parnas, A., Mortensen, E. L., & Parnas, J. (2012). Premorbid intelligence and educational level in bipolar and unipolar disorders: A Danish draft board study. *Journal of Affective Disorders*, *136*(3), 1188–1191. https://doi.org/10.1016/j.jad.2011.12.007

Sparding, T., Joas, E., Clements, C., Sellgren, C. M., Pålsson, E., & Landén, M. (2021). Long-term trajectory of cognitive performance in people with bipolar disorder and controls: 6-year longitudinal study. *BJPsych Open*, *7*(e115), 1–7. https://doi.org/10.1192/bjo.2021.66

Sparding, T., Silander, K., Pålsson, E., Östlind, J., Sellgren, C., Ekman, C. J., Joas, E., Hansen, S., & Landén, M. (2015). Cognitive functioning in clinically stable patients with bipolar disorder I and II. *PLoS ONE*, *10*(1), e0115562. https://doi.org/10.1371/journal.pone.0115562

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rücker, G., Harbord, R. M., Schmid, C. H., Tetzlaff, J., Deeks, J. J., Peters, J., Macaskill, P., Schwarzer, G., Duval, S., Altman, D. G., Moher, D., & Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ (Online)*, *343*(7818), 1–8. https://doi.org/10.1136/bmj.d4002

Strawbridge, R., Tsapekos, D., Hodsoll, J., Mantingh, T., Yalin, N., McCrone, P., Boadu, J., Macritchie, K., Cella, M., Reeder, C., Fish, J., Wykes, T., & Young, A. H. (2021). Cognitive remediation therapy for patients with bipolar disorder: A randomised proof-of-concept trial. *Bipolar Disorders*, *23*(2), 196–208. https://doi.org/10.1111/bdi.12968

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/H0054651

Sun, N., Yang, C., He, X., Liu, Z., Liu, S., Li, X., Wang, Y., Jin, R., & Zhang, K. (2020). Impact of expression and genetic variation of microrna-34b/c on cognitive dysfunction in patients with major depressive disorder. *Neuropsychiatric Disease and Treatment*, *16*, 1543–1554. https://doi.org/10.2147/NDT.S247787

Sweeney, J. A., Kmiec, 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. https://doi.org/10.1016/S0006-3223(00)00910-0

Talarowska, M., Florkowski, A., Zboralski, K., Berent, D., Wierzbiński, P., & Gałecki, P. (2010). Auditory-verbal declarative and operating memory among patients suffering from depressive

disorders - preliminary study. *Advances in Medical Sciences*, *55*(2), 317–327. https://doi.org/10.2478/v10039-010-0053-0

Tavares, J. V. T., 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. https://doi.org/10.1016/j.biopsych.2007.05.034

The Cochrane Collaboration. (2014). *Review Manager (RevMan)* (5.3). The Nordic Cochrane Centre.

Thompson, J. M., Gallagher, P., Hughes, J. H., Watson, S., Gray, J. M., Ferrier, I. N., & Young, A. H. (2005). Neurocognitive impairment in euthymic patients with bipolar affective disorder. *British Journal of Psychiatry*, *186*(JAN.), 32–40. https://doi.org/10.1192/bjp.186.1.32

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. https://doi.org/10.1037/a0014740

Torrent, C., Martinez-Arán, A., Daban, C., Amann, B., Balanzá-Martínez, V., Del Mar Bonnín, C., Cruz, N., Franco, C., Tabarés-Seisdedos, R., & Vieta, E. (2011). Effects of atypical antipsychotics on neurocognition in euthymic bipolar patients. *Comprehensive Psychiatry*, *52*(6), 613–622. https://doi.org/10.1016/j.comppsych.2010.12.009

Torres, I. J., Boudreau, V. G., & Yatham, L. N. (2007). Neuropsychological functioning in euthymic bipolar disorder: A meta-analysis. *Acta Psychiatrica Scandinavica*, *116*(SUPPL. 434), 17–26. https://doi.org/10.1111/j.1600-0447.2007.01055.x

Trivedi, J. K., Goel, D., Sharma, S., Singh, A. P., Sinha, P. K., & Tandon, R. (2007). Cognitive functions in stable schizophrenia and euthymic state of bipolar disorder. *Indian Journal of Medical Research*, *126*(5), 433–439.

Tsitsipa, E., & Fountoulakis, K. N. (2015). The neurocognitive functioning in bipolar disorder: a systematic review of data. *Annals of General Psychiatry*, *14*, 42. https://doi.org/10.1186/s12991-015-0081-z

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, *45*(3).

van den Heuvel, M. P., & Sporns, O. (2019). A cross-disorder connectome landscape of brain dysconnectivity. *Nature Reviews Neuroscience*, *1*. https://doi.org/10.1038/s41583-019-0177-6

Van Rheenen, T. E., Lewandowski, K. E., Bauer, I. E., Kapczinski, F., Miskowiak, K. W., Burdick, K. E., & Balanzá-Martínez, V. (2020). Current understandings of the trajectory and emerging correlates of cognitive impairment in bipolar disorder: An overview of evidence. *Bipolar Disorders*, *22*(1), 13–27. https://doi.org/10.1111/bdi.12821

Van Rheenen, T. E., Lewandowski, K. E., Tan, E. J., Ospina, L. H., Ongur, D., Neill, E., Gurvich, C., Pantelis, C., Malhotra, A. K., Rossell, S. L., & Burdick, K. E. (2017). Characterizing cognitive heterogeneity on the schizophrenia-bipolar disorder spectrum. *Psychological Medicine*, *47*(10), 1848–1864. https://doi.org/10.1017/S0033291717000307

van Velzen, L. S., Kelly, S., Isaev, D., Aleman, A., Aftanas, L. I., Bauer, J., Baune, B. T., Brak, I. V., Carballedo, A., Connolly, C. G., Couvy-Duchesne, B., Cullen, K. R., Danilenko, K. V., Dannlowski, U., Enneking, V., Filimonova, E., Förster, K., Frodl, T., Gotlib, I. H., … Schmaal, L. (2020). White matter disturbances in major depressive disorder: a coordinated analysis across 20 international cohorts in the ENIGMA MDD working group. *Molecular Psychiatry*, *25*(7), 1511–1525. https://doi.org/10.1038/s41380-019-0477-2

Vaskinn, A., Sundet, K., Simonsen, C., Hellvin, T., Melle, I., & Andreassen, O. A. (2011). Sex Differences in Neuropsychological Performance and Social Functioning in Schizophrenia and Bipolar

Disorder. *Neuropsychology*, *25*(4), 499–510. https://doi.org/10.1037/a0022677

Vicent-Gil, M., Keymer-Gausset, A., Serra-Blasco, M., Carceller-Sindreu, M., de Diego-Adeliño, J., Trujols, J., Mur, M., Pérez, V., Alvarez, E., Cardoner, N., & Portella, M. J. (2018). Cognitive predictors of illness course at 12 months after first-episode of depression. *European Neuropsychopharmacology*, *28*(4), 529–537. https://doi.org/10.1016/j.euroneuro.2018.02.001

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Walhovd, K. B., Krogsrud, S. K., Amlien, I. K., Bartsch, H., Bjørnerud, A., Due-Tønnessen, P., Grydeland, H., Hagler, D. J., Håberg, A. K., Kremen, W. S., Ferschmann, L., Nyberg, L., Panizzon, M. S., Rohani, D. A., Skranes, J., Storsve, A. B., Sølsnes, A. E., Tamnes, C. K., Thompson, W. K., … Fjell, A. M. (2016). Neurodevelopmental origins of lifespan changes in brain and cognition. *PNAS*, *113*(33), 9357–9362. https://doi.org/10.1073/pnas.1524259113

Walsh, N. D., Seal, M. L., Williams, S. C. R., & Mehta, M. A. (2009). An investigation of cognitive "branching" processes in major depression. *BMC Psychiatry*, *9*, 69. https://doi.org/10.1186/1471-244X-9-69

Wang, H. T., Smallwood, J., Mourao-Miranda, J., Xia, C. H., Satterthwaite, T. D., Bassett, D. S., & Bzdok, D. (2020). Finding the needle in a high-dimensional haystack: Canonical correlation analysis for neuroscientists. *NeuroImage*, *216*, 116745. https://doi.org/10.1016/j.neuroimage.2020.116745

Wang, M., Chen, W., Liu, B., Ju, Y., Dong, Q., Lu, X., Sun, J., Zhang, L., Guo, H., Zhao, F., Li, W., Zhang, L., Li, Z., Liao, M., Zhang, Y., Liu, J., & Li, L. (2022). Sleep disturbances and retardation in the prediction of cognitive impairments in patients with major depressive disorder. *Research Square Preprint*. https://doi.org/10.21203/rs.3.rs-2174102/v1

Wang, T., Wang, K., Qu, H., Zhou, J., Li, Q., Deng, Z., Du, X., Lv, F., Ren, G., Guo, J., Qiu, J., & Xie, P. (2016). Disorganized cortical thickness covariance network in major depressive disorder implicated by aberrant hubs in large-scale networks. *Scientific Reports*, *6*, 27964. https://doi.org/10.1038/srep27964

Wang, X., Zhou, H., & Zhu, X. (2020). Attention deficits in adults with Major depressive disorder: A systematic review and meta-analysis. *Asian Journal of Psychiatry*, *53*, 102359. https://doi.org/10.1016/j.ajp.2020.102359

Wang, Y., Leiberg, K., Ludwig, T., Little, B., Necus, J. H., Winston, G., Vos, S. B., Tisi, J. de, Duncan, J. S., Taylor, P. N., & Mota, B. (2021). Independent components of human brain morphology. *NeuroImage*, *226*, 117546. https://doi.org/10.1016/j.neuroimage.2020.117546

Wang, Y., Necus, J., Kaiser, M., & Mota, B. (2016). Universality in human cortical folding in health and disease. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(45), 12820–12825. https://doi.org/10.1073/pnas.1610175113

Wang, Y., Zhong, S., Jia, Y., Zhou, Z., Wang, B., Pan, J., & Huang, L. (2015). Interhemispheric resting state functional connectivity abnormalities in unipolar depression and bipolar depression. *Bipolar Disorders*, *17*(5), 486–495. https://doi.org/10.1111/bdi.12315

Wardenaar, K. J., & de Jonge, P. (2013). Diagnostic heterogeneity in psychiatry: towards an empirical solution. *BMC Medicine*, *11*, 201. https://doi.org/10.1186/1741-7015-11-201

Watson, S., Gallagher, P., Dougall, D., Porter, R., Moncrieff, J., Ferrier, I. N., & Young, A. H. (2014). Childhood trauma in bipolar disorder. *Australian and New Zealand Journal of Psychiatry*, *48*(6), 564–570. https://doi.org/10.1177/0004867413516681

Watson, S., Gallagher, P., Porter, R. J., Smith, M. S., Herron, L. J., Bulmer, S., Young, A. H., & Ferrier, I. N. (2012). A randomized trial to examine the effect of mifepristone on neuropsychological

performance and mood in patients with bipolar depression. *Biological Psychiatry*, *72*(11), 943–949. https://doi.org/10.1016/j.biopsych.2012.05.029

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence - Second Edition (WASI-II)*. NCS Pearson.

Weise, C. M., Bachmann, T., Schroeter, M. L., & Saur, D. (2019). When less is more: Structural correlates of core executive functions in young adults – A VBM and cortical thickness study. *NeuroImage*, *189*(September 2018), 896–903. https://doi.org/10.1016/j.neuroimage.2019.01.070

Weissman, A. S., & Bates, M. E. (2010). Increased clinical and neurocognitive impairment in children with autism spectrum disorders and comorbid bipolar disorder. *Research in Autism Spectrum Disorders*, *4*(4), 670–680. https://doi.org/10.1016/j.rasd.2010.01.005

Weissman, D. H., Roberts, K. C., Visscher, K. M., & Woldorff, M. G. (2006). The neural bases of momentary lapses in attention. *Nature Neuroscience*, *9*(7), 971–978. https://doi.org/10.1038/nn1727

Weshsler, D. (1997). *Wechsler Adult Intelligence Scale* (3rd ed.). Psychological Corporation.

Westlye, L. T., Grydeland, H., Walhovd, K. B., & Fjell, A. M. (2011). Associations between regional cortical thickness and attentional networks as measured by the attention network test. *Cerebral Cortex*, *21*(2), 345–356. https://doi.org/10.1093/cercor/bhq101

Whelan, R. (2008). Effective Analysis of Reaction Time Data. *The Psychological Record*, *58*, 475–482.

White, D. A., Myerson, J., & Hale, S. (1997). How cognitive is psychomotor slowing in depression? Evidence from a meta-analysis. *Aging, Neuropsychology, and Cognition*, *4*(3), 166–174. https://doi.org/10.1080/13825589708256645

Wise, T., Radua, J., Nortje, G., Cleare, A. J., Young, A. H., & Arnone, D. (2016). Voxel-based meta-analytical evidence of structural disconnectivity in major depression and bipolar disorder. *Biological Psychiatry*, *79*(4), 293–302. https://doi.org/10.1016/j.biopsych.2015.03.004

Withall, A., Harris, L. M., & Cumming, S. R. (2009). The relationship between cognitive function and clinical and functional outcomes in major depressive disorder. *Psychological Medicine*, *39*(3), 393–402. https://doi.org/10.1017/S0033291708003620

World Health Organization. (1992). The ICD-10 Classification of Mental and Behavioural Disorders. *International Classification*, *10*, 1–267. https://doi.org/10.1002/1520-6505(2000)9:5<201::AID-EVAN2>3.3.CO;2-P

Wu, M.-J., Mwangi, B., Bauer, I. E., Passos, I. C., Sanches, M., Zunta-Soares, G. B., Meyer, T. D., Hasan, K. M., & Soares, J. C. (2017). Identification and individualized prediction of clinical phenotypes in bipolar disorders using neurocognitive data, neuroimaging scans and machine learning. *NeuroImage*, *145*, 254–264. https://doi.org/10.1016/j.neuroimage.2016.02.016

Xu, G., Lin, K., Rao, D., Dang, Y., Ouyang, H., Guo, Y., Ma, J., & Chen, J. (2012). Neuropsychological performance in bipolar I, bipolar II and unipolar depression patients: A longitudinal, naturalistic study. *Journal of Affective Disorders*, *136*(3), 328–339. https://doi.org/10.1016/j.jad.2011.11.029

Yadav, S. K., Alreja, S., Sengar, K. S., & Singh, A. R. (2011). Cognitive flexibility in patients with bipolar affective disorder: Current episode mania. *Journal of Projective Psychology & Mental Health*, *18*(2), 200–205.

Yamamoto, T., & Shimada, H. (2012). Cognitive dysfunctions after recovery from major depressive episodes. *Applied Neuropsychology*, *19*(3), 183–191. https://doi.org/10.1080/09084282.2011.643959

Yatham, L. N., Torres, I. J., Malhi, G. S., Frangou, S., Glahn, D. C., Bearden, C. E., Burdick, K. E., Martínez-Arán, A., Dittmann, S., Goldberg, J. F., Ozerdem, A., Aydemir, O., & Chengappa, K. N. R. (2010). The International Society for Bipolar Disorders-Battery for Assessment of Neurocognition (ISBD-BANC). *Bipolar Disorders*, *12*(4), 351–363. https://doi.org/10.1111/j.1399-5618.2010.00830.x

Young, A. H., Rigney, U., Shaw, S., Emmas, C., & Thompson, J. M. (2011). Annual cost of managing bipolar disorder to the UK healthcare system. *Journal of Affective Disorders*, *133*(3), 450–456. https://doi.org/10.1016/j.jad.2011.06.016

Young, R. C., Biggs, J. T., Ziegler, V. E., & Meyer, D. A. (1978). A rating scale for mania: Reliability, validity and sensitivity. *British Journal of Psychiatry*, *133*(11), 429–435. https://doi.org/10.1192/bjp.133.5.429

Yu, H., Li, M. L., Meng, Y., Li, X. J., Wei, W., Li, Y. F., Li, L., Guo, W., Wang, Q., Deng, W., Ma, X., Coid, J., & Li, T. (2021). Inferior frontal gyrus seed-based resting-state functional connectivity and sustained attention across manic/hypomanic, euthymic and depressive phases of bipolar disorder. *Journal of Affective Disorders*, *282*, 930–938. https://doi.org/10.1016/j.jad.2020.12.199

Zaremba, D., Schulze Kalthoff, I., Förster, K., Redlich, R., Grotegerd, D., Leehr, E. J., Meinert, S., Dohm, K., Bürger, C., Enneking, V., Böhnlein, J., Repple, J., Opel, N., Jörgens, S., Yüksel, D., Schmitt, S., Stein, F., Kircher, T., Krug, A., … Dannlowski, U. (2019). The effects of processing speed on memory impairment in patients with major depressive disorder. *Progress in Neuropsychopharmacology and Biological Psychiatry*, *92*, 494–500. https://doi.org/10.1016/j.pnpbp.2019.02.015

Zarp Petersen, J., Varo, C., Skovsen, C. F., Ott, C. V., Kjærstad, H. L., Vieta, E., Harmer, C. J., Knudsen, G. M., Kessing, L. V., Macoveanu, J., & Miskowiak, K. W. (2022). Neuronal underpinnings of cognitive impairment in bipolar disorder: A large data-driven functional magnetic resonance imaging study. *Bipolar Disorders*, *24*(1), 69–81. https://doi.org/10.1111/bdi.13100

Zazula, R., Mohebbi, M., Dodd, S., Dean, O. M., Berk, M., Vargas, H. O., & Nunes, S. O. V. (2021). Cognitive Profile and Relationship with Quality of Life and Psychosocial Functioning in Mood Disorders. *Archives of Clinical Neuropsychology*, *37*(2), 376–389. https://doi.org/10.1093/arclin/acab054

Zhang, B. H., Feng, L., Feng, Y., Xin, L. M., Zhu, X. Y., Tan, Y. L., & Wang, G. (2020). The Effect of Cognitive Impairment on the Prognosis of Major Depressive Disorder. *The Journal of Nervous and Mental Disease*, *208*(9), 683–688. https://doi.org/10.1097/NMD.0000000000001180

Zhang, W., Zhu, N., Lai, J., Liu, J., H Ng, C., Chen, J., Qian, C., Du, Y., Hu, C., Chen, J., Hu, J., Wang, Z., Zhou, H., Xu, Y., Fang, Y., Shi, C., & Hu, S. (2020). Reliability and validity of thinc-it in evaluating cognitive function of patients with bipolar depression. *Neuropsychiatric Disease and Treatment*, *16*, 2419–2428. https://doi.org/10.2147/NDT.S266642

Zhao, S., Wang, X., Chen, Z., Zhou, H., Han, Y., Tang, H., Yan, R., Yao, Z., & Lu, Q. (2021). Sex differences in the association between symptom profiles and cognitive functioning in patients with depressive disorder. *Journal of Affective Disorders*, *287*, 1–7. https://doi.org/10.1016/j.jad.2021.03.020

Zhou, Y., Zheng, W., Liu, W., Wang, C., Zhan, Y., Li, H., Chen, L., & Ning, Y. (2019). Cross-sectional relationship between kynurenine pathway metabolites and cognitive function in major depressive disorder. *Psychoneuroendocrinology*, *101*, 72–79. https://doi.org/10.1016/j.psyneuen.2018.11.001

Zhu, Y., Womer, F. Y., Leng, H., Chang, M., Yin, Z., Wei, Y., Zhou, Q., Fu, S., Deng, X., Lv, J., Song, Y., Ma, Y., Sun, X., Bao, J., Wei, S., Jiang, X., Tan, S., Tang, Y., & Wang, F. (2019). The relationship between cognitive dysfunction and symptom dimensions across schizophrenia, bipolar disorder,

and major depressive disorder. *Frontiers in Psychiatry*, *10*(APR), 1–8. https://doi.org/10.3389/fpsyt.2019.00253

Zhu, Z., Zhao, Y., Wen, K., Li, Q., Pan, N., Fu, S., Li, F., Radua, J., Vieta, E., Kemp, G. J., Biswa, B. B., & Gong, Q. (2022). Cortical thickness abnormalities in patients with bipolar disorder: A systematic review and meta-analysis. *Journal of Affective Disorders*, *300*, 209–218. https://doi.org/10.1016/j.jad.2021.12.080

Zimmerman, M. E., Brickman, A. M., Paul, R. H., Grieve, S. M., Tate, D. F., Gunstad, J., Cohen, R. A., Aloia, M. S., Williams, L. M., Clark, C. R., Whitford, T. J., & Gordon, E. (2006). The relationship between frontal gray matter volume and cognition varies across the healthy adult lifespan. *American Journal of Geriatric Psychiatry*, *14*(10), 823–833. https://doi.org/10.1097/01.JGP.0000238502.40963.ac