Characterisation of Deficits and Sex Differences in Verbal and Visual Memory/Learning in Bipolar Disorder

Andrea Gogos[1] , Jinwon Son[2], Susan L. Rossell[3], James Karantonis[3,4], Lisa S. Furlong[4] , Kim Felmingham[2] and Tamsyn E. Van Rheenen[3,4,*]

[1]The Florey Institute of Neuroscience and Mental Health, Parkville, VIC, Australia
[2]School of Psychological Sciences, The University of Melbourne, Parkville, VIC, Australia
[3]Centre for Mental Health, School of Health Sciences, Faculty of Health, Arts and Design, Swinburne University, Melbourne, VIC, Australia
[4]Melbourne Neuropsychiatry Centre, Department of Psychiatry, The University of Melbourne, Parkville, VIC, Australia

Abstract

Objective: Cognitive impairment is consistently reported in bipolar disorder (BD), but few studies have characterised which memory component processes are affected. Further, it is unknown whether the component processes underlying memory impairment are moderated by sex. The present study examined diagnosis and sex differences in both verbal and visual memory/learning domains in patients with BD and psychiatrically healthy controls. **Method:** Verbal and visual memory/learning were measured using the Hopkins Verbal Learning Test-Revised (HVLT-R) and Brief Visuospatial Memory Test-Revised (BVMT-R). 114 patients with BD ($n = 50$ males, $n = 64$ females), were compared to 105 psychiatrically healthy controls ($n = 42$ males, $n = 63$ females). **Results:** Patients with BD had worse performance in verbal and visual immediate and total recall, verbal and visual delayed free recall, and verbal recognition discrimination scores, but there were no group differences in learning slopes and cumulative learning index scores. There were trends for BD females to outperform BD males in visual memory/learning free recall and cumulative learning, but these results did not survive multiple testing correction. These findings did not change in a secondary sensitivity analysis comparing only strictly euthymic BD patients to controls ($n = 64$). **Conclusion:** The present study found trait-like verbal and visual memory/learning impairment in BD that was attributable to deficient encoding and/or consolidation processes rather than deficits in learning. We did not find marked sex differences in either visual or verbal memory/learning measures, although some trend level effects were apparent and deserve exploration in future studies.

Keywords: HVLT-R, BVMT-R, Encoding, Consolidation, Learning, Male, Female

INTRODUCTION

Neurocognitive impairment is a major predictor of functional disability in bipolar disorder (BD), with 40–60% of patients showing symptoms of cognitive dysfunction (Burdick et al., 2014; Martinez-Aran et al., 2004; Sole et al., 2017; Van Rheenen et al., 2017). Several meta-analyses have highlighted that neurocognitive impairments in BD are significant in the areas of memory/learning, attention, processing speed, executive function, response inhibition and set shifting; and are present during both symptomatic and euthymic illness phases (Arts, Jabben, Krabbendam & van Os, 2008; Bora, Yucel & Pantelis, 2009; Robinson et al., 2006; Torres, Boudreau & Yatham, 2007). One of the most consistently

reported deficits in BD is in the domain of memory/learning, which encompasses encoding (initial acquisition of information into memory), consolidation (information storage) and retrieval (remembering) of information. Specifically, we and others found that compared to healthy control participants, patients with BD have significantly worse performance in tasks measuring learning, immediate memory/attention (Altshuler et al., 2004; Bearden et al., 2006; Gogos, Joshua & Rossell, 2010; Nitzburg et al., 2017; van Gorp, Altshuler, Theberge, Wilkins & Dixon, 1998; Van Rheenen & Rossell, 2014b) and delayed recall and recognition (Deckersbach, Savage et al., 2004; Nitzburg et al., 2004; van Gorp et al., 1998; Van Rheenen & Rossell, 2014b). Impairment in verbal memory in particular, may represent an endophenotype for BD given its presence in unaffected relatives with the disorder (Calafiore, Rossell & Van Rheenen, 2018). It is also one of the most strongly predictive of all

---

*Correspondence and reprint requests to: Dr Tamsyn E. Van Rheenen, Melbourne Neuropsychiatry Centre, Level 3, Alan Gilbert Building, 161 Barry St, Carlton, VIC 3053, Australia. E-mail: tamsyn.van@unimelb.edu.au

common cognitive domains in terms of poor occupational, social and global functioning outcomes (Atre-Vaidya et al., 1998; Martinez-Aran et al., 2004).

Currently, there is scarce literature examining sex differences in memory/learning in BD (Gogos, Ney, Seymour, Van Rheenen & Felmingham, 2019), despite reported sex differences in BD in terms of epidemiology and clinical presentation of the illness (Bucker et al., 2014; Gogos et al., 2019; Gogos, van den Buuse & Rossell, 2009). In healthy individuals there is evidence of a female advantage for verbal memory, and some evidence of a male advantage for visuospatial memory (Kramer, Delis & Daniel, 1988; Pauls, Petermann & Lepach, 2013). However, in BD, two existing studies found that patients had lower verbal memory/learning performance compared to healthy control participants, but effects of sex were not evident (Bucker et al., 2004; Gogos et al., 2010). By contrast, another study showed that male patients with BD had worse verbal memory/learning performance compared to male controls, but this deficit was not observed when comparing female patients with BD to female controls (Carrus et al., 2010). These studies focused on single measures of memory/learning but did not analyse the influence of sex on the component processes involved. Yet understanding the specific fundamental processes that contribute to memory deficits in BD is important for knowing where to focus cognitive treatments that might aim to remediate these impairments.

In previous work we have characterised a range of verbal memory/learning component processes in both male and female patients with BD together (Van Rheenen & Rossell, 2014b), finding that verbal memory impairments in BD are attributable to deficient encoding processes during the acquisition phase of learning (Van Rheenen & Rossell, 2014b). However, in that work we did not investigate whether similar cognitive mechanisms underpin visual memory/learning performance in BD, nor did we examine the influence of sex explicitly.

In comparison to verbal memory/learning, impairment in visual memory/learning is less explored in BD. Some studies of BD show that patients have impaired immediate recall on visual memory/learning tasks compared to controls, which may relate to an encoding rather than a retention deficit (Deckersbach, McMurrich et al., 2004). By contrast, other studies have found no group differences in visual memory/learning performance (Altshuler et al., 2004; van Gorp et al., 1998). In the few studies examining sex differences on this domain, Carrus et al. (2004) found that male patients with BD had worse performance than male controls in immediate, but not delayed, visual memory/learning, and there were no differences between female patients with BD and controls (Carrus et al., 2004). Tournikioti et al., (2018) found that male controls outperformed female controls in visuospatial recognition memory, but this sex difference was not observed in patients with BD (Tournikioti et al., 2004). By contrast, we found no sex differences in visual memory/learning performance (Gogos et al., 2010). Of note, each of these studies used a different measure of visual memory/learning

and had a relatively small sample size ($n = 40$; Gogos et al., 2010). Such heterogeneity across a limited number of studies makes comparison of results difficult. However, these previous studies suggest there may be some sex differences, and this warrants further investigation, particularly in the context of visual memory component processes which have not been explored in BD in the context of sex effects so far.

In this exploratory study we aimed to examine verbal and visual memory performance in BD, with a focus on identifying the extent to which there are sex differences in these domains compared to controls. In doing so, we aimed to address some of the methodological limitations of previous studies. Firstly, we employed a larger sample size than previous work to overcome issues with low statistical power. Secondly, we comprehensively measured the different components of verbal and visual memory/learning (such as delayed memory, retention rate, and learning slope) using well recognised tasks. We did this to define the profile of memory performance in BD more accurately and identify the extent to which different component processes are affected by sex, since this remains unknown due to the paucity of literature on the topic. We hypothesised that patients with BD would be more impaired than healthy control participants on measures of verbal and visual memory/learning performance. We also hypothesised that females would outperform males in verbal memory/learning irrespective of diagnosis, given strong past evidence indicating this effect in the general population (Kramer et al., 1988; Pauls et al., 2013). Given sparse literature examining sex differences in cognition in BD itself, we did not make explicit hypotheses regarding sex by diagnosis interaction effects for memory/learning.

METHODS

Participants

Data from 219 participants (92 males, 127 females) was obtained from a databank held by the senior author. All participants had participated in studies led by the senior author and as such, the memory/learning performance of parts of the sample have been reported previously (Karantonis et al., 2020; Van Rheenen & Rossell, 2014a, 2014b), albeit not in the context of sex differences. All participants had been recruited using general advertisements as well as online websites and social media; patients with BD were also recruited through community support groups (exclusion criteria detailed below). All participants had given prior informed consent for the analysis of stored data. Participants were selected for the current analysis from the databank if responses on the primary study measures - the Hopkins Verbal Learning Test-Revised (HVLT-R: Brandt & Benedict, 2001) and/or the Brief Visuospatial Memory Test-Revised (BVMT-R: Benedict, 2007), were recorded. That is, participants in the databank with missing data on the primary study measures were excluded from analyses.

All procedures contributing to this work complied with the ethical standards of the relevant Human Ethics Review Boards and with the Declaration of Helsinki.

The clinical group comprised $n = 114$ patients ($n = 50$ males, $n = 64$ females) with a DSM-IV-TR diagnosis [confirmed using the Mini International Neuropsychiatric Interview (MINI: Sheehan et al., 1998)] of BD (BD-I $n = 101$; BD-II $n = 13$). A total of $n = 34$ BD participants reported a family history of BD or a relative treated with a mood stabilizer and $n = 30$ a history of rapid cycling; $n = 60$ BD participants had a current comorbid anxiety disorder and $n = 15$ had current comorbid substance *abuse*. No participants met criteria for substance-induced BD or an organic medical cause. Current mood symptom severity was assessed using the Young Mania Rating Scale (YMRS: Young, Biggs, Ziegler & Meyer, 1978) and the Montgomery-Asberg Depression Rating Scale (MADRS: Montgomery & Asberg, 1979), from which 64 patients with BD (56%) were considered to be strictly euthymic, as defined by MADRS and YMRS scores of <8. The remaining 50 (44%) of the BD sample displayed mild-moderate symptoms (i.e. 31 with MADRS scores > 8, 6 with YMRS scores > 8, and 13 with both MADRS and YMRS scores > 8). Age of diagnosis was self-reported by the patients with BD. Eighteen patients with BD were unmedicated and 96 were medicated at the time of assessment ($n = 77$ mood stabilisers, $n = 56$ antipsychotics, $n = 36$ antidepressants, $n = 13$ other psychotropic drugs).

A control sample of 105 psychiatrically healthy participants ($n = 42$ males, $n = 63$ females) were recruited by general advertisement. The MINI screen was performed to confirm that no one in the control group had a current diagnosis of psychiatric illnesses. Participants were explicitly asked at phone screening and in the demographic questionnaire, and then subsequently excluded from the control group, if they reported having ever been diagnosed (lifetime) with any type of psychiatric illness.

Exclusion criteria for participants included: not within the 18–65 year age range, difficulties with spoken English, a history of traumatic brain injury, hearing or visual impairments, neurological or degenerative illness, alcohol or substance *dependence* in the past 3 months, pregnancy, a history of psychotropic medication use (healthy controls only), a family history (first or second degree relative) of a mood or psychiatric disorder (healthy controls only). See Tables 1 and 2; and Supplementary Tables 1–3 for full demographic details.

Cognitive Measures

Verbal memory/learning was measured with the HVLT-R (Brandt & Benedict, 2001) and visual memory/learning was measured with the BVMT-R (Benedict, 2007). The HVLT-R and BVMT-R were designed to be methodologically similar to each other, with both tests having good validity (Benedict, Schretlen, Groninger, Dobraski & Shpritz, 1996; Shapiro, Benedict, Schretlen & Brandt, 1999) and test-retest reliability (Benedict, Schretlen, Groninger &

Brandt, 1998; Nuechterlein et al., 2008). These tests were given to participants as part of broader cognitive batteries, of which the reporting is beyond the scope of the current study. Other cognitive tests that were given to participants between the learning and delayed trials of the HVLT-R and BVMT-R (see below for more detail) varied from person-to-person depending on the study in which they were originally enrolled (see Karantonis et al. (2004); Van Rheenen & Rossell (2013a, 2013b, 2014a, 2014b, 2014c) for further detail on some of these tests). Thus, these tests were deemed unlikely to substantively influence participant performance.

On the HVLT-R, participants were required to verbally recall twelve target words over three learning trials (T1–T3), and a delayed recall trial (T4) 20–30 min later. The task also included a recognition trial comprising 24 words (12 target words from the original list, 12 unrelated distractor words of which 6 were semantically related), that were presented directly following the delayed trial. The HVLT-R was scored for the following variables of interest: (a) free recall (number of correctly recalled words on each of T1, T2, and T3); (b) total recall (sum of T1–T3); (c) learning slope (difference between T3 and T1; Benedict et al., 1998); (d) cumulative word learning (learning slope multiplied by total recall score; Foster et al., 2009); (e) delayed free recall (T4); (f) retention percentage rate (percentage of T4 divided by the higher of T2 or T3); (g) recognition discrimination index (total number of true positives minus the total number of false positives on the recognition trial). Raw scores were used for analyses.

On the BVMT-R, participants were required to visually recall six geometric shapes over three learning trials (T1–T3). A delayed trial (T4) was completed 20–30 min later. A score of 2 was given for any accurately reproduced and placed geometric shapes. Any accurately drawn but incorrectly placed or any inaccurately drawn but recognizable reproduced shapes were scored a 1. Finally, absent or unrecognizable and incorrectly placed redrawing were scored a 0. The highest possible score for each learning trial was 12. The BVMT-R was scored for the following variables of interest: (a) free recall (number of correctly recalled shapes on each of T1, T2, and T3); (b) total recall (sum of T1–T3); (c) learning slope (difference between T3 and T1); (d) cumulative shape learning (learning slope multiplied by total recall score); (e) delayed free recall (T4); (f) retention percentage rate (percentage rate of T4 divided by the higher of T2 or T3). Raw scores were used for analyses.

The initial words/shapes recalled on T1 were taken as a measure of attention, and the total recall score as the amount of information learnt across each trial. The learning slope reflected the degree of learning across the learning trials (i.e. the average number of newly recalled words/shapes over the learning trials), whereas cumulative word/shape learning reflected the interaction between the learning slope and the number of words/shapes recalled across these trials (i.e. it captured both the degree of learning and the amount of information learned across trials). Retention rate in turn, reflected the amount of information retained between the immediate

Table 1. Demographic characteristics of the participants

	BD				HC				Comparisons					
	Male (n = 50)		Female (n = 64)		Male (n = 42)		Female (n = 63)		Dx		Sex		Dx × Sex	
	M	SD	M	SD	M	SD	M	SD	F	p	F	p	F	p
Age (years)	42.5	11.73	35.6	11.83	34.1	14.33	33.3	13.25	7.84	**0.006**	3.98	**0.047**	2.58	0.110
Education (years)	17.5	5.21	16.2	4.05	15.9	3.56	16.5	4.21	0.86	0.355	0.29	0.592	2.04	0.155

Note: BD = bipolar disorder; HC = healthy control; Dx = diagnosis; M = mean; SD = standard deviation.
Bolded values are significant at $p < 0.05$.

and delayed learning trials. The recognition discrimination index (verbal memory only) reflected the capacity to discriminate between previously presented (words presented on initial trials) and new information (new distractor words presented in recognition trial).

Statistical Analysis

All statistical analyses were performed using the Statistical Package for Social Sciences (version 25.0; SPSS, Chicago, IL). The data conformed to the assumptions of the relevant statistical tests prior to analyses. Univariate analyses of variance (ANOVA) and chi-square tests were conducted to examine diagnostic differences and/or interactions on demographic and clinical variables. A diagnostic and sex effect were detected for age, thus subsequent ANOVAs for the measures of interest were conducted with and without age included as a covariate. The inclusion of age had no effect and therefore only the results of analyses without covarying for age are presented below. Further, although there were no diagnosis, sex or diagnosis by sex interaction effects for years of education, given the potential of this variable to influence cognitive performance we also correlated years of education with each of the measures of interest in the full sample in unreported preliminary analyses. As no correlations were significant even at an uncorrected p-value of $p < 0.05$, in the absence of diagnostic and sex effects/interactions we reasoned that years of education is unlikely to have a substantial effect on the cognitive findings and thus we did not control for years of education in the subsequent ANOVAs conducted with the measures of interest as outcomes. In these ANOVAs, diagnosis and sex were specified as fixed factors, and each of the measures of interest as dependent variables. First, multivariate ANOVAs including all verbal and visual memory domains were performed separately to obtain omnibus diagnosis, sex, and diagnosis by sex interaction effects. Second, univariate ANOVAs were then conducted for individual verbal memory/learning scores of interest. These results were corrected for multiple testing using a false discovery rate (FDR) of 5% ($p < 0.05$) for the main effects of diagnosis, sex, and the group by sex interaction separately. Trend level effects were taken as those significant at $p < 0.05$ but not surviving FDR correction. It should be noted that we conducted a sensitivity power analysis (Lakens, 2021) using G*power

3.1.9.4 for ANOVA ($\alpha = 0.05$, power = 0.8, total sample = 219, number of groups = 2) to ascertain power in our sample, from which we determined that $f = 0.19$ (small to medium) was the lowest effect size detectable.

Given the presence of some symptomatic patients in the BD sample, in a secondary sensitivity analysis these procedures were conducted again including only euthymic BD patients to examine trait-effects.[1] Further bivariate Pearson's correlations between variables of interest and MADRS and YMRS scores were also examined in the full BD sample. Given varied medication use in the BD sample, further exploratory analyses comparing patients on and off different classes of medication (mood stabilisers, antipsychotics and antidepressants) were also conducted using multivariate and univariate ANOVAs with follow-up univariate ANOVAs per sex for significant interactions. There was some sporadic missing data for a small number of participants on some variables (see Tables), but data for the delayed recall trial of the BVMT-R (T4) was missing in $n = 50$ BD and $n = 75$ controls, thus caution should be taken when interpreting the results of the analysis of this variable and BVMT-R retention (which is derived from T4). Where relevant, effect sizes are reported as Cohen's *d*.

RESULTS

Demographic and Clinical Characteristics

There was no significant difference between BD and control participants in terms of sex ($\chi^2 = 0.36$, $p = 0.55$) and education (Table 1, Supplementary Table 1). There were main effects of sex and diagnosis for age, with males being older than females and patients with BD being older than controls (Table 1, Supplementary Table 1). There were no significant differences between male and female patients with BD for any of the clinical variables measured (i.e. MADRS, YMRS, age of diagnosis, symptom status, BD subtype, psychosis history, current comorbid anxiety or substance use, family history of BD, rapid cycling, medication use and type; Table 2).

---

[1] We also re-analysed the data excluding the 13 BD II participants in the sample, to identify if diagnostic subtype had an effect. Given this resulted in a minor drop in the size of the sample, and as there were no meaningful differences compared to the results of the full sample analysis, the results of this subsample analysis are not presented for brevity.

Table 2. Clinical characteristics of the BD patients

	BD male ($n = 50$)			BD female ($n = 64$)			Comparisons	
	n	M	SD	n	M	SD	F/χ^2	p
MADRS		9.0	8.28		9.9	9.12	0.27	0.604
YMRS		4.5	3.81		4.9	4.77	0.28	0.595
Age of diagnosis		30.1	10.79		27.4	10.14	1.70	0.195
Symptom status (euthymic/symptomatic)	28/22			36/28			0.03	0.985
Subtype (BD I/II)	47/3			54/10			2.09	0.148
Psychosis history (yes/no)[a]	33/16			35/26			1.14	0.285
Comorbid anxiety disorder (yes/no)	26/24			34/30			0.01	0.905
Comorbid substance/alcohol abuse (yes/no)	6/44			9/55			0.11	0.746
Family history of BD or mood stabilisers (yes/no)[b]	18/31			16/45			1.40	0.236
Rapid cycling (yes/no)[c]	16/34			14/48			1.25	0.263
Mood stabilisers (yes/no)	34/16			43/21			0.01	0.927
Antipsychotics (yes/no)	23/27			33/31			0.35	0.556
Antidepressants (yes/no)	13/37			23/41			1.28	0.257
Other psychiatric meds (yes/no)	7/43			6/58			0.59	0.441
No meds (yes/no)	8/42			10/54			0.00	0.957

Notes: BD = bipolar disorder; MADRS = Montgomery-Asberg Depression Rating Scale; M = mean; meds = medication use; SD = standard deviation; YMRS = Young Mania Rating Scale.
[a] Data missing for $n = 4$ participants.
[b] Data missing for one male and three female participants.
[c] Data missing for two female participants.

Primary Analysis of Verbal and Visual Memory/Learning in All Participants

When comparing all BD and control participants, there were omnibus diagnostic effects for both verbal (F(9,186) = 4.98, $p < 0.001$) and visual (F(7, 83) = 2.404, $p = 0.027$) memory/learning domains, but no omnibus sex or diagnosis by sex interaction effects were evident (Table 3).

For individual verbal memory/learning scores there were main effects of diagnosis for the following variables (Table 3): HVLT-R free recall T1 ($d = 0.68$), T2 ($d = 0.73$), T3 ($d = 0.78$), total recall ($d = 0.81$), delayed memory ($d = 0.87$), retention ($d = 0.60$), and recognition discrimination ($d = 0.62$); all BD < controls (Figure 1, Supplementary Table 5). There were main effects of sex for HVLT-R free recall T1 ($d = 0.31$, male < female) but this did not survive correction (see also Supplementary Table 6). There were no other main effects of sex or diagnosis by sex interactions (Supplementary Table 4).

For individual visual memory/learning scores there were main effects of diagnosis for BVMT-R free recall T1 ($d = 0.66$), T2 ($d = 0.78$), T3 ($d = 0.65$), total recall ($d = 0.78$) and delayed memory ($d = 0.66$), all surviving correction (all BD < controls; Table 3, Supplementary Table 5; Figure 1). There were main effects of sex for BVMT-R free recall T1 and T2 ($d = 0.37$, $d = 0.28$; females > males; Supplementary Table 6), and diagnosis by sex interactions for BVMT-R free recall T3 (BD males < BD females < control females < control males) and cumulative shape learning (control females < BD males < BD females < control males), but none of these effects survived correction (Table 3, Supplementary Table 4). No other main effects or interactions were evident.

There were no significant correlations between any of the variables of interest (verbal and visual memory/learning) and MADRS or YMRS scores (data not shown).

Secondary Sensitivity Analysis of Verbal and Visual Memory/Learning in Controls and Only the Euthymic BD Patients

When comparing euthymic BD patients with control participants, there were omnibus diagnostic effects for verbal memory/learning (F(9,138) = 4.98, $p < 0.001$) but no omnibus diagnostic effects for visual memory/learning. No omnibus sex or diagnosis by sex interaction effects were evident for either verbal or visual memory/learning (Supplementary Table 7).

As per the analysis including all participants, for individual verbal memory/learning scores there were significant main effects of diagnosis for the following variables: HVLT-R free recall T1 ($d = 0.64$), T2 ($d = 0.83$), T3 ($d = 0.83$), total recall ($d = 0.83$), delayed memory ($d = 1.03$), retention ($d = 0.79$), and recognition discrimination ($d = 0.73$), all surviving FDR correction (all euthymic BD < controls). There were no significant main effects of sex or diagnosis by sex interactions (Supplementary Tables 7–10).

For individual visual memory/learning scores there were significant main effects of diagnosis for BVMT-R free recall T1 ($d = 0.66$), T2 ($d = 0.78$), T3 ($d = 0.78$), total recall ($d = 0.82$) and visual delayed memory ($d = 0.58$), all surviving correction (all euthymic BD < controls). No main effects of sex or diagnosis by sex interactions were evident (Supplementary Tables 7–10).

Table 3. Descriptives, main effects and interactions for the variables of interest

	BD (n = 114)						HC (n = 105)						Effect					
	Male			Female			Male			Female			Dx		Sex		Dx × Sex	
	M	SD	n	M	SD	n	M	SD	n	M	SD	n	F	p	F	p	F	p
Verbal memory/learning																		
Omnibus													4.984	<0.001[b]	0.671	0.734	0.787	0.628
Free recall T1	6.42	2.01	50	6.89	1.54	64	7.55	1.77	42	8.13	1.80	62	23.53	**<0.001**	4.65	0.032[a]	0.05	0.821
Free recall T2	8.68	1.93	50	8.97	1.98	64	10.17	1.75	42	10.16	1.54	62	29.08	**<0.001**	0.33	0.569	0.35	0.554
Free recall T3	9.58	1.96	50	9.77	1.61	64	10.93	1.40	42	10.90	1.31	62	32.55	**<0.001**	0.14	0.713	0.23	0.629
Total recall (T1–T3 sum)	24.68	5.15	50	25.63	4.17	64	28.69	4.18	42	28.95	4.44	63	35.46	**<0.001**	0.96	0.329	0.31	0.580
Delayed recall (T4)	8.22	2.53	50	8.64	2.03	64	10.12	1.87	34	10.37	1.72	51	36.36	**<0.001**	1.26	0.263	0.76	0.783
Learning slope	3.34	1.66	50	3.16	1.52	64	3.50	1.64	42	2.87	1.54	62	0.08	0.774	3.48	0.063	1.05	0.308
CWL	80.47	45.05	50	79.91	43.64	64	99.19	45.56	42	80.19	42.65	62	2.22	0.146	2.45	0.119	1.97	0.162
Retention (%)	83.22	17.91	50	85.63	15.10	64	91.94	12.27	34	94.40	12.09	51	16.65	**<0.001**	1.29	0.258	0.00	0.991
Recognition DI	10.16	1.75	49	10.48	1.65	64	11.09	1.17	35	11.35	0.91	51	18.25	**<0.001**	1.97	0.162	0.02	0.900
Visual memory/learning																		
Omnibus													2.404	0.027[b]	0.539	0.803	0.170	0.991
Free recall T1	5.24	2.62	50	6.38	2.92	64	7.21	2.70	42	8.02	2.55	61	25.54	**<0.001**	6.76	0.010[a]	0.20	0.656
Free recall T2	8.00	2.93	50	9.14	2.68	64	10.40	1.58	42	10.54	1.85	61	34.44	**<0.001**	3.88	0.050[a]	2.40	0.123
Free recall T3	9.50	2.57	50	10.53	2.11	64	11.40	1.27	42	11.25	1.14	61	27.77	**<0.001**	2.86	0.092	5.32	0.022[a]
Total recall (T1–T3 sum)	22.74	7.46	50	26.05	6.92	64	29.02	4.71	42	29.80	5.02	63	35.46	**<0.001**	0.96	0.329	0.31	0.580
Delayed recall (T4)	9.32	2.34	34	10.29	2.21	31	10.92	1.61	13	11.31	1.01	16	8.22	**0.005**	2.20	0.142	0.40	0.529
Learning slope	4.34	1.90	50	4.23	2.34	64	4.29	2.53	42	3.30	2.04	61	2.68	0.103	3.27	0.072	2.13	0.146
CSL	98.10	49.01	50	104.41	57.85	64	116.93	61.41	41	91.15	51.60	61	0.13	0.714	1.65	0.201	4.47	0.036[a]
Retention (%)	96.23	14.49	34	99.18	14.09	31	95.98	12.31	13	98.45	5.95	16	0.03	0.866	0.86	0.356	0.01	0.935

Note: BD = bipolar disorder; HC = healthy control; CSL = cumulative shape learning; CWL = cumulative word learning; DI = discrimination index; Dx = diagnosis; M = mean; SD = standard deviation; T1 = trial 1; T2 = trial 2; T3 = trial 3; T4 = trial 4.
Verbal memory/learning was assessed with the Hopkins Verbal Learning Test – Revised, and visual memory/learning was measured with the Brief Visuospatial Memory Test – Revised.
Higher scores indicate better performance.
All *p*-values represent uncorrected *p*-values.
Bolded values survived FDR correction ($p < 0.05$).
[a] Are significant at $p < 0.05$ uncorrected but did not survive FDR correction.
[b] Omnibus test is significant at $p < 0.05$ uncorrected (no correction applied).
Note that Cohen's *d* effect sizes are reported in the supplementary material.

Exploratory Medication Analysis in the BD Sample

Mood stabilisers

There were no main effects of sex or mood stabiliser use, or sex by mood stabiliser use interaction effects at either the omnibus or individual test score level for either verbal or visual memory/learning (Supplementary Table 11).

Antipsychotics

For verbal memory/learning there were no omnibus effects of sex or antipsychotic use, but an omnibus sex by antipsychotic use interaction effect was evident ($F(8, 102) = 2.11$, $p = 0.041$). There was an antipsychotic use by sex interaction effect for HVLT-R free recall T3, and main effects of antipsychotic use for HVLT-R free recall (T1, T2 and T3) and total recall (not using > using).

For visual memory/learning there were no omnibus effects of sex or antipsychotic use, nor a sex by antipsychotic use interaction. However, there were main effects of sex (females > males) for BVMT-R free recall (T1, T2 and T3), total recall and delayed recall. None of these effects survived FDR correction (Supplementary Table 12).

Antidepressants

For verbal memory/learning there were no omnibus effects of sex or antidepressant use, nor a sex by antidepressant use interaction. There was a main effect of sex for HVLT-R free recall T1 and total recall (females > males) and a main effect of antidepressant use for recognition discrimination index (using > not using), however neither survived FDR correction. Sex by antidepressant use interaction effects were also evident for the HVLT-R free recall (T1, T2 and T3), total
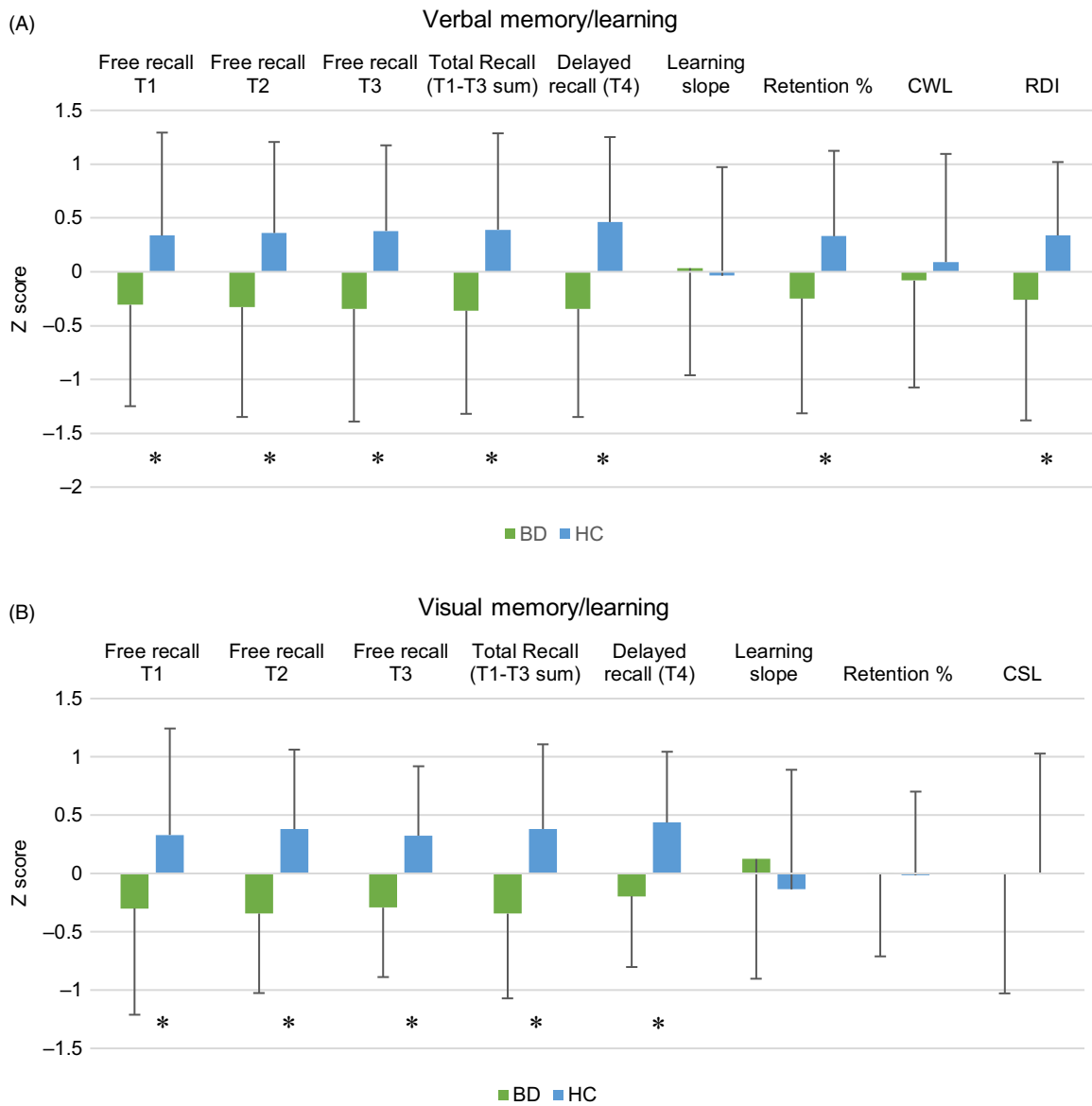
Fig. 1. Mean ± SD for main effects of diagnosis for the variables of interest. Graph (A) represents verbal memory/learning performance measured with the Hopkins Verbal Learning Test-Revised (HVLT-R); Graph (B) represents visual memory/learning performance measured with the Brief Visuospatial Memory Test-Revised (BVMT-R). Verbal and visual memory/learning scores were standardised against the total sample mean (BD and control) for scaling and visualisation purposes. BD = bipolar disorder; HC = healthy control; CSL = cumulative shape learning; CWL = cumulative word learning; T1 = trial 1; T2 = trial 2; T3 = trial 3; T4 = trial 4; RDI = recognition discrimination index. *$p < 0.05$ FDR corrected.

recall, delayed recall, and recognition discrimination index (Supplementary Table 13). Follow up analysis for the interactions that survived FDR correction (Supplementary Table 14) showed that for HVLT-R T2 scores, males had significantly higher scores if not using rather than using antidepressants ($d = 0.60$), and females tended to have better scores if using compared to not using antidepressants ($d = 0.52$). For HVLT-R T3 and total recall scores, females had significantly higher scores if using than not using antidepressants ($d = 0.69$ and $d = 0.72$), but there was no difference for males ($d = 0.15$ and $d = 0.55$).

For visual memory/learning there were no omnibus effects of sex or antidepressant use, nor a sex by antidepressant use interaction. There were significant main effects of sex for BVMT-R scores when patients were separated by antidepressant use (Supplementary Table 13); females outperformed males on BVMT-R free recall (T1, T2 and T3), total recall and delayed recall. There was also a sex by antidepressant use interaction for BVMT-R free recall T2, total recall and delayed recall (females had higher scores if using while males had higher scores if not using), however these interactions did not survive FDR correction.

DISCUSSION

It is clearly established that a large proportion of patients with BD exhibit cognitive deficits (Van Rheenen et al., 2020). However, far less is known about the extent of memory impairment in BD and whether sex differences in memory performance exist. As such, this study aimed to present new knowledge on the underlying component processes involved in verbal and visual memory/learning in BD and whether these are moderated by sex. In support of our first hypothesis, the main findings were that patients with BD had worse performance than healthy controls measured by omnibus diagnosis effects as well as individual domain measures of verbal and visual memory/learning. However, our hypothesis concerning sex effects was not supported because although there were some trend level (at $p < 0.05$ uncorrected) sex differences and interaction effects for some measures of verbal and visual memory/learning performance, none survived FDR correction.

The present data support the majority of studies showing that patients with BD have verbal (Altshuler et al., 2004; Bearden et al., 2004; Deckersbach, Savage et al., 2004; Gogos et al., 2010; Nitzburg et al., 2004; van Gorp et al., 1998; Van Rheenen & Rossell, 2014b) and visual memory/learning (Carrus et al., 2004; Deckersbach, McMurrich et al., 2004) deficits. Similar to our previous study using a subset of this sample (Van Rheenen & Rossell, 2014b) as well as others (Carrus et al., 2004; Nitzburg et al., 2004) investigating the component processes of verbal memory/learning in BD, the present study found that patients with BD had worse performance than healthy controls in terms of verbal and visual total recall (an index of learning), as well as verbal and visual delayed free recall (an index of information storage); all with medium-large effect sizes. Relevantly, patients with BD had greater difficulty in the immediate free recall of the HVLT-R word list and BVMT-R shape list compared to healthy controls. This deficit was apparent from trial 1, suggesting that attention is also deficient in BD.

The recognition discrimination index was only measured for verbal memory/learning, where patients with BD also performed worse than healthy controls. However, there were no differences between the groups in terms of their verbal and visual learning slopes (rate of learning), or cumulative learning index which takes into account the interaction between the total number of words/shapes recalled and the average number of newly recalled correct words/shapes across the learning trials (i.e. the slope). Thus, it appears that the verbal and visual learning impairments seen in patients with BD were primarily reflective of a reduced capacity to encode information on each trial, rather than in the *degree* to which new information is learnt across them.

Between-group differences in both verbal and visual memory/learning also emerged in terms of the proportion of information retained after a delay of 20–30 minutes, where effect sizes and significance levels indicated that patients with BD had worse verbal, but not visual retention percentage scores compared to controls. Deficient verbal retention scores

in patients with BD suggest a more rapid loss of information learned previously, which may reflect difficulties in the consolidation (storage) of verbal information. However, in this case, recognition discrimination scores would also be expected to be equally or more severely impaired in BD than delayed recall scores. In our data they were not (e.g. effect sizes were $d = 0.87$ for delayed recall but $d = 0.60$ for recognition). Thus, it is possible that the reduced retention scores seen in patients with BD here were also due to difficulties in the retrieval of verbal information after a delay. Indeed, the absence of a retention impairment in BD for visual information, in combination with impaired performance on trials 1–3, suggests that patients with BD in this sample have a deficit in their capacity to encode visual information during the acquisition phase of learning but not in their ability to retain the visual information they do encode. The large BD deficit in visual delayed recall may thus be due to an issue with information retrieval rather than consolidation/storage. However, it should be noted that we had a smaller sample size for the delayed recall/retention rate scores for visual memory/learning, and no recognition trials were presented for this test. As a result of these limitations, this result and interpretation should be considered with caution. Nonetheless, impaired scores on memory tests, particularly those indexing immediate and delayed recall, but also recognition, have been linked to mild cognitive impairment and Alzheimer's disease (Weissberger et al., 2017). Since an association between BD and dementia has already been established (Kessing, Olsen, Mortensen & Andersen, 1999), it is possible that the impaired memory scores seen here may index risk for the development of this and related conditions in BD. Further research is required to explore this further.

In terms of sex effects, the results showed that there were no significant overall sex differences in any of the components of these measures, nor were there significant diagnosis by sex interactions. Although previous literature on sex differences in memory/learning in BD is scarce, the current findings support some (Bucker et al., 2004; Gogos et al., 2010) but not all existing research (Carrus et al., 2004; Tournikioti et al., 2004). Relevantly, the sex difference in visual memory/learning identified in a study by Tournikioti et al. (2004) was due to a difference between male and female controls rather than a sex difference in patients with BD. Further, the discrepancy of the findings of this study and that of Carrus et al. (2004) may be due to differences in the cognitive measures used. Indeed, Carrus et al. (2004) used the California Verbal Learning Test, which comprises more words and more trials than the HVLT-R used here. It may thus be more sensitive than the HVLT-R in BD if sex effects are subtle. The limitations of the HVLT-R are particularly relevant to interpreting our finding of no sex differences between male and female healthy controls, as the scores on T2, T3 and delayed recall trials of the HVLT-R were high for these participants, indicating potential ceiling effects. Hence, future studies would do well to examine sex differences in BD using more sensitive memory/learning measures to see if this leads to different outcomes.

Other possible reasons for the discrepancies between studies relate to differences in sample characteristics, such as sample size, clinical (e.g. number of mood episodes) and developmental (e.g., childhood trauma or socio-economic status) history, BD subtype, age or mood state. Our study was limited in that we did not have objective mood episode or developmental history data available. Although our sample included significantly older patients with BD compared to controls, further analyses covarying for age found no effect of age on the results. In addition, analyses excluding the 13 patients with BD II in the sample indicated that diagnostic subtype had no effect on the pattern of findings in this data. While our sample included patients with BD in both euthymic and symptomatic mood states, current mood status did not affect the results as analyses comparing euthymic BD patients and controls revealed the same pattern of results as when symptomatic BD patients were included. This was supported by the lack of correlations between the variables of interest and mood assessment scores. These findings correspond with a large body of research indicating that memory impairment is a stable trait-like feature of BD (Bearden et al., 2004; Bora et al., 2009; Robinson et al., 2004; Van Rheenen & Rossell, 2014b).

Sex differences in memory/learning in the general population have been examined extensively, with several studies showing that females outperform males on measures of verbal memory/learning (Beatty, Mold & Gontkovsky, 2003; Gogos, 2013; Gogos et al., 2010; Kramer et al., 1988). Although no sex effects survived FDR correction, the present study did find a trend for an overall effect of sex for verbal and visual free recall (trial 1; immediate memory/attention); with females outperforming males. In addition, there were trend-level diagnosis by sex effects for visual memory/learning free recall (trial 3), and cumulative shape learning, where BD females outperformed BD males. This trend for sex differences may be due to the presence of certain sex hormone levels (Gogos et al., 2019; Luine, 2014; Sbisa, van den Buuse & Gogos, 2017). Indeed, data from our previous study suggested that sex hormone status of healthy female participants may be the determining factor on whether a sex difference in cognition is observed (Gogos, 2013). For example, we found a sex difference in cognition where women with high levels of sex hormones (either natural or synthetic) outperformed men, however women with low levels of sex hormones and males performed similarly (Gogos, 2013). Of the five studies (including the present study) examining sex differences in memory/learning (Bucker et al., 2004; Carrus et al., 2004; Gogos, 2013; Tournikioti et al., 2004), only Gogos et al., (2010) controlled for fluctuations in sex hormones in women (by only testing women during times of low levels of circulating sex hormones). Future studies should examine the role of sex hormones in mediating cognitive performance in patients with BD.

There is some evidence that mood stabilizing medications in patients with BD can either positively or negatively affect different cognitive domains, including memory (Pachet &

Wisniewski, 2003; Sabater et al., 2016); although these studies did not examine sex differences. For example, Sabater et al. (2004) found that those treated with lithium monotherapy had preserved verbal memory/learning, while patients with BD on anticonvulsant monotherapy showed impairments in visual memory/learning. By contrast, Tournikioti et al. (2017) suggested that sodium valproate predicted better performance on a visual memory/learning task in patients with BD (Tournikioti et al., 2004). In our study, comparisons of patients on and off different classes of medications revealed no main effects of sex or sex by medication interactions for mood stabilisers. There was trend level evidence that antipsychotic use impaired verbal learning performance in terms of each learning trial and total recall irrespective of sex, however this should be interpreted with caution as the effect did not survive correction for multiple comparisons. There was also evidence that males not using antidepressants had better verbal learning performance on trial 2 of the HVLT-R than those using them, with an opposite effect evident in females. Females using antidepressants also had better verbal learning trial 3 and total recall (learning) scores than females not using them; an effect not evident in males. However, given the small number of male patients using antidepressants, these results should be viewed with caution. Further, since patients in our sample were largely polytherapy users who were treated with combinations of different classes of drugs, we were unable to dissect the specific effects of different psychotropic medications on memory/learning. Future research should investigate sex differences in cognition in medication-free or monotherapy patients with BD. This is particularly relevant given some evidence that lithium treatment indirectly effects oestrogen signalling via its effects of glycogen synthase kinase-3 (Gould & Manji, 2005). Although we excluded participants that had neurological or degenerative illnesses, we did not have information about other medical conditions, such as cardiovascular disease or diabetes, which are commonly comorbid with BD and have been linked to cognitive impairment (Van Rheenen, McIntyre, Balanza-Martinez, Berk & Rossell, 2021). Thus, future studies should carefully examine the influence of comorbid medical conditions and their treatments as they may alter cognitive performance.

Previous research on sex differences in BD has primarily focused on the epidemiology and clinical presentation of the illness. Given that cognitive impairment is evident in BD, the present study sought to investigate whether sex differences exist in verbal and visual memory/learning. We did not find marked sex differences in either measure of memory/learning, although some trend level effects were apparent and deserve exploration in future studies. Relevantly, the present findings supported previous work by showing a verbal memory/learning impairment in BD that is attributable to deficient encoding and consolidation processes. Further, it extended the literature by indicating a similar pattern of results in visual memory/learning highlighting deficient encoding processes. These deficits were apparent

independent of mood status, diagnostic subtype and age, supporting the notion that cognitive dysfunction represents a trait-like feature of the illness.

ACKNOWLEDGEMENTS

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S1355617721001442

CONFLICTS OF INTEREST

The authors have nothing to disclose.

References

Altshuler L.L., Ventura J., van Gorp W.G., Green M.F., Theberge D.C., & Mintz J. (2004). Neurocognitive function in clinically stable men with bipolar I disorder or schizophrenia and normal control subjects. *Biological Psychiatry*, 56(8), 560–569. doi: 10.1016/j.biopsych.2004.08.002

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. doi: 10.1017/S0033291707001675

Atre-Vaidya N., Taylor M.A., Seidenberg M., Reed R., Perrine A., & Glick-Oberwise F. (1998). Cognitive deficits, psychopathology, and psychosocial functioning in bipolar mood disorder. *Neuropsychiatry Neuropsychology and Behavioral Neurology*, 11(3), 120–126.

Bearden C.E., Glahn D.C., Monkul E.S., Barrett J., Najt P., Kaur S., & Soares J.C. (2006). Sources of declarative memory impairment in bipolar disorder: Mnemonic processes and clinical features. *Journal of Psychiatric Research*, 40(1), 47–58. doi: 10.1016/j.jpsychires.2005.08.006

Beatty W.W., Mold J.W., & Gontkovsky S.T. (2003). RBANS performance: Influences of sex and education. *J Clin Exp Neuropsychol*, 25(8), 1065–1069. doi: 10.1076/jcen.25.8.1065.16732

Benedict R. (2007). *Brief Visuospatial Memory Test-Revised professional manual*. Odessa, FL: Psychological Assessment Resources.

Benedict R., Schretlen D., Groninger L., & Brandt J. (1998). Hopkins Verbal Learning Test – Revised: Normative data and analysis of inter-form and test-retest reliability. *The Clinical Neuropsychologist*, 12(1), 43–55. doi: 10.1076/clin.12.1.43.1726

Benedict R., Schretlen D., Groninger L., Dobraski M., & Shpritz B. (1996). Revision of the Brief Visuospatial Memory Test: Studies on normal performance, reliability, and validity. *Psychological Assessment*, 8(2), 145–153. doi: 10.1037/1040-3590.8.2.145, doi:https://doi.org/10.1037/1040-3590.8.2.145

Bora E., Yucel M., & Pantelis C. (2009). Cognitive endophenotypes of bipolar disorder: A meta-analysis of neuropsychological deficits in euthymic patients and their first-degree relatives. *Journal of Affective Disorders*, 113(1-2), 1–20. doi: 10.1016/j.jad.2008.06.009

Brandt J., & Benedict R. (2001). *Hopkins Verbal Learning Test Revised: Professional manual*. Lutz, FL: Psychological Assessment Resources.

Bucker J., Popuri S., Muralidharan K., Kozicky J.M., Baitz H.A., Honer W.G., & Yatham L.N. (2014). Sex differences in cognitive functioning in patients with bipolar disorder who recently recovered from a first episode of mania: Data from the Systematic Treatment Optimization Program for Early Mania (STOP-EM). *Journal of Affective Disorders*, 155, 162–168. doi: 10.1016/j.jad.2013.10.044

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. doi: 10.1017/S0033291714000439

Calafiore D., Rossell S.L., & Van Rheenen T.E. (2018). Cognitive abilities in first-degree relatives of individuals with bipolar disorder. *Journal of Affective Disorders*, 225, 147–152. doi: 10.1016/j.jad.2017.08.029

Carrus D., Christodoulou T., Hadjulis M., Haldane M., Galea A., Koukopoulos A., & Frangou S. (2010). Gender differences in immediate memory in bipolar disorder. *Psychological Medicine*, 40(8), 1349–1355. doi: 10.1017/S0033291709991644

Deckersbach T., McMurrich S., Ogutha J., Savage C.R., Sachs G., & Rauch S.L. (2004). Characteristics of non-verbal memory impairment in bipolar disorder: The role of encoding strategies. *Psychological Medicine*, 34(5), 823–832. doi: 10.1017/s0033291703001685

Deckersbach T., Savage C.R., Reilly-Harrington N., Clark L., Sachs G., & Rauch S.L. (2004). Episodic memory impairment in bipolar disorder and obsessive-compulsive disorder: The role of memory strategies. *Bipolar Disorders*, 6(3), 233–244. doi: 10.1111/j.1399-5618.2004.00118.x

Foster P.S., Drago V., Crucian G.P., Rhodes R.D., Shenal B.V., & Heilman K.M. (2009). Verbal learning in Alzheimer's disease: Cumulative word knowledge gains across learning trials. *Journal of the International Neuropsychological Society*, 15(5), 730–739. doi: 10.1017/s1355617709990336

Gogos A. (2013). Natural and synthetic sex hormones: Effects on higher-order cognitive function and prepulse inhibition. *Biological Psychology*, 93(1), 17–23. doi: 10.1016/j.biopsycho.2013.02.001

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. doi: 10.3109/00048670903446882

Gogos A., Ney L.J., Seymour N., Van Rheenen T.E., & Felmingham K.L. (2019). Sex differences in schizophrenia, bipolar disorder, and post-traumatic stress disorder: Are gonadal hormones the link? *British Journal of Pharmacology*, *176*(21), 4119–4135. doi: 10.1111/bph.14584

Gogos A., van den Buuse M., & Rossell S. (2009). Gender differences in prepulse inhibition (PPI) in bipolar disorder: Men have reduced PPI, women have increased PPI. *International Journal of Neuropsychopharmacology*, *12*(9), 1249–1259. doi: 10.1017/S1461145709000480

Gould T.D., & Manji H.K. (2005). Glycogen synthase kinase-3: A putative molecular target for lithium mimetic drugs. *Neuropsychopharmacology*, *30*(7), 1223–1237. doi: 10.1038/sj.npp.1300731

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

Kessing L.V., Olsen E.W., Mortensen P.B., & Andersen P.K. (1999). Dementia in affective disorder: A case-register study. *Acta Psychiatrica Scandinavica*, *100*(3), 176–185. doi: 10.1111/j.1600-0447.1999.tb10843.x

Kramer J.H., Delis D.C., & Daniel M. (1988). Sex differences in verbal learning. *Journal of Clinical Psychology*, *44*(6), 907–915.

Lakens D. (2021). *Sample size justification*. APA, 10.31234/osf.io/9d3yf.

Luine V.N. (2014). Estradiol and cognitive function: Past, present and future. *Hormones and Behavior*, *66*(4), 602–618. doi: 10.1016/j.yhbeh.2014.08.011

Martinez-Aran A., Vieta E., Colom F., Torrent C., Sanchez-Moreno J., Reinares M., & Salamero M. (2004). Cognitive impairment in euthymic bipolar patients: Implications for clinical and functional outcome. *Bipolar Disorders*, *6*(3), 224–232. doi: 10.1111/j.1399-5618.2004.00111.x

Montgomery S.A., & Asberg M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, *134*, 382–389. doi: 10.1192/bjp.134.4.382

Nitzburg G.C., Cuesta-Diaz A., Ospina L.H., Russo M., Shanahan M., Perez-Rodriguez M., & Burdick K.E. (2017). Organizational learning strategies and verbal memory deficits in bipolar disorder. *Journal of the International Neuropsychological Society*, *23*(4), 358–366. doi: 10.1017/S1355617717000133

Nuechterlein K.H., Green M.F., Kern R.S., Baade L.E., Barch D.M., Cohen J.D., & Marder S.R. (2008). The MATRICS consensus cognitive battery, part 1: Test selection, reliability, and validity. *American Journal of Psychiatry*, *165*(2), 203–213. doi: 10.1176/appi.ajp.2007.07010042

Pachet A.K., & Wisniewski A.M. (2003). The effects of lithium on cognition: An updated review. *Psychopharmacology (Berlin)*, *170*(3), 225–234. doi: 10.1007/s00213-003-1592-x

Pauls F., Petermann F., & Lepach A.C. (2013). Gender differences in episodic memory and visual working memory including the effects of age. *Memory*, *21*(7), 857–874.

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*(1-3), 105–115. doi: 10.1016/j.jad.2006.02.016

Sabater A., Garcia-Blanco A.C., Verdet H.M., Sierra P., Ribes J., Villar I., & Livianos L. (2016). Comparative neurocognitive effects of lithium and anticonvulsants in long-term stable bipolar patients. *Journal of Affective Disorders*, *190*, 34–40. doi: 10.1016/j.jad.2015.10.008

Sbisa A.M., van den Buuse M., & Gogos A. (2017). The effect of estradiol and its analogues on cognition in preclinical and clinical research: Relevance to schizophrenia. In P.A. Gargiulo & H.L. Mesones-Arroyo (Eds.), *Psychiatry and neuroscience update: Bridging the divide* (Vol. 2). Springer.

Shapiro A.M., Benedict R.H., Schretlen D., & Brandt J. (1999). Construct and concurrent validity of the Hopkins Verbal Learning Test-revised. *Clinical Neuropsychologist*, *13*(3), 348–358. doi: 10.1076/clin.13.3.348.1749

Sheehan D.V., Lecrubier Y., Sheehan K.H., Amorim P., Janavs J., Weiller E., & Dunbar G.C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry*, *59 Suppl 20*, 22–33, quiz 34–57.

Sole B., Jimenez E., Torrent C., Reinares M., Bonnin C.D.M., Torres I., & Vieta E. (2017). Cognitive impairment in bipolar disorder: Treatment and prevention strategies. *International Journal of Neuropsychopharmacology*, *20*(8), 670–680. doi: 10.1093/ijnp/pyx032

Torres I.J., Boudreau V.G., & Yatham L.N. (2007). Neuropsychological functioning in euthymic bipolar disorder: A meta-analysis. *Acta Psychiatrica Scandinavica*, *434*, 17–26. doi: 10.1111/j.1600-0447.2007.01055.x,

Tournikioti K., Ferentinos P., Michopoulos I., Alevizaki M., Soldatos C.R., Dikeos D., & Douzenis A. (2017). Clinical and treatment-related predictors of cognition in bipolar disorder: Focus on visual paired associative learning. *European Archives of Psychiatry and Clinical Neuroscience*, *267*(7), 661–669. doi: 10.1007/s00406-016-0743-0

Tournikioti K., Ferentinos P., Michopoulos I., Dikeos D., Soldatos C.R., & Douzenis A. (2018). Sex-related variation of neurocognitive functioning in bipolar disorder: Focus on visual memory and associative learning. *Psychiatry Research*, *267*, 499–505. doi: 10.1016/j.psychres.2018.06.037

van Gorp W.G., Altshuler L., Theberge D.C., Wilkins J., & Dixon W. (1998). Cognitive impairment in euthymic bipolar patients with and without prior alcohol dependence. A preliminary study. *Archives of General Psychiatry*, *55*(1), 41–46. doi: 10.1001/archpsyc.55.1.41

Van Rheenen T.E., Lewandowski K.E., Bauer I.E., Kapczinski F., Miskowiak K., Burdick K.E., & Balanza-Martinez 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. doi: 10.1111/bdi.12821

Van Rheenen T.E., Lewandowski K.E., Tan E.J., Ospina L.H., Ongur D., Neill E., & Burdick K.E. (2017). Characterizing cognitive heterogeneity on the schizophrenia-bipolar disorder spectrum. *Psychological Medicine*, *47*(10), 1848–1864. doi: 10.1017/S0033291717000307

Van Rheenen T.E., McIntyre R.S., Balanza-Martinez V., Berk M., & Rossell S.L. (2021). Cumulative cardiovascular disease risk and triglycerides differentially relate to subdomains of executive function in bipolar disorder; preliminary findings. *Journal of Affective Disorders*, *278*, 556–562. doi: 10.1016/j.jad.2020.09.104

Van Rheenen T.E., & Rossell S.L. (2013a). Auditory-prosodic processing in bipolar disorder; from sensory perception to

emotion. *Journal of Affective Disorders*, *151*(3), 1102–1107. doi: 10.1016/j.jad.2013.08.039

Van Rheenen T.E., & Rossell S.L. (2013b). Picture sequencing task performance indicates theory of mind deficit in bipolar disorder. *Journal of Affective Disorders*, *151*(3), 1132–1134. doi: 10.1016/j.jad.2013.07.009

Van Rheenen T.E., & Rossell S.L. (2014a). An empirical evaluation of the MATRICS Consensus Cognitive Battery in bipolar disorder. *Bipolar Disorders*, *16*(3), 318–325. doi: 10.1111/bdi.12134

Van Rheenen T.E., & Rossell S.L. (2014b). Investigation of the component processes involved in verbal declarative memory function in bipolar disorder: Utility of the Hopkins Verbal Learning Test-Revised. *Journal of the International Neuropsychological Society*, *20*(7), 727–735. doi: 10.1017/S1355617714000484

Van Rheenen T.E., & Rossell S.L. (2014c). Multimodal emotion integration in bipolar disorder: An investigation of involuntary cross-modal influences between facial and prosodic channels. *Journal of the International Neuropsychological Society*, *20*(5), 525–533. doi: 10.1017/S1355617714000253

Weissberger G.H., Strong J.V., Stefanidis K.B., Summers M.J., Bondi M.W., & Stricker N.H. (2017). Diagnostic accuracy of memory measures in Alzheimer's dementia and mild cognitive impairment: A systematic review and meta-analysis. *Neuropsychology Review*, *27*(4), 354–388. doi: 10.1007/s11065-017-9360-6

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*, 429–435. doi: 10.1192/bjp.133.5.429