TECHNICAL COMMENTARY

NeuRA Discover. Conquer. Cure.
BIPOLAR DISORDERS LIBRARY

Learning

Introduction

Learning is the ability to acquire, or change, existing knowledge, behaviours or skills. There are two distinct forms of learning: explicit (or declarative) learning occurs during a high level of consciousness regarding specific learnt content, for example, memorising information for an exam. Implicit (or procedural) learning is less conscious and refers to learning that is gained from task performance, for example, juggling. Explicit verbal learning can be measured with the Hopkins Verbal Learning test, the California Verbal Learning test and verbal list-learning, for example. The Brief Visuospatial memory test, the Rey design learning test, the Rey complex figure test, and visual reproduction all measure explicit visual learning. Implicit learning can be measured using the Serial Reaction Time task where learning is inferred from reduced reaction time to stimuli.

Method

We have included only systematic reviews (systematic literature search, detailed methodology with inclusion/exclusion criteria) published in full text, in English, from the year 2010 that report results separately for people with a diagnosis of bipolar and related disorders. Reviews were identified by searching the databases MEDLINE, EMBASE, and PsycINFO. Hand searching reference lists of identified reviews was also conducted. When multiple copies of review topics were found, only the most recent and comprehensive review was included. Reviews with pooled data are prioritised for inclusion.

Review reporting assessment was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist that describes a preferred way to present a meta-analysis[1]. Reviews with less than 50% of items checked have been excluded from the library. The PRISMA flow diagram is a suggested way of providing information about studies included and excluded with reasons for exclusion. Where no flow diagram has been presented by individual reviews, but identified studies have been described in the text, reviews have been checked for this item. Note that early reviews may have been guided by less stringent reporting checklists than the PRISMA, and that some reviews may have been limited by journal guidelines.

Evidence was graded using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group approach where high quality evidence such as that gained from randomised controlled trials (RCTs) may be downgraded to moderate or low if review and study quality is limited, if there is inconsistency in results, indirect comparisons, imprecise or sparse data and high probability of reporting bias. It may also be downgraded if risks associated with the intervention or other matter under review are high. Conversely, low quality evidence such as that gained from observational studies may be upgraded if effect sizes are large or if there is a dose dependent response. We have also taken into account sample size and whether results are consistent, precise and direct with low associated risks (see end of table for an explanation of these terms)[2]. The resulting table represents an objective summary of the available evidence, although the conclusions are solely the opinion of staff of NeuRA (Neuroscience Research Australia).

Results

We found nine systematic reviews that met our inclusion criteria[3-11].

- High quality evidence suggests medium-sized effects of poorer verbal and visual learning in people with bipolar disorder, (including first-episode) compared to controls. There was also a small trend effect in young first-degree relatives of people with bipolar disorder (aged 10-25 years).

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Learning

- Moderate to high quality evidence suggests a medium-sized effect of better list learning in people with first-episode bipolar disorder compared to people with first-episode schizophrenia.
- Moderate quality evidence suggests a large, significant effect of poorer verbal and visual learning and memory in euthymic youth with bipolar disorder (aged 13 years) compared to age and IQ matched controls.
- Moderate to low quality evidence also suggests a medium-sized effect of poorer serial learning in elderly people with bipolar disorder compared to age and education matched controls.
- High quality evidence suggests a medium-sized effect of poorer list learning in people with bipolar disorder and a history of psychotic symptoms compared to people with bipolar disorder without a history of psychotic symptoms. There was a small effect of poorer list learning in people with bipolar I disorder compared to people with bipolar II disorder.
- Moderate quality evidence suggests a medium-sized effect of better list learning in people with major depression than people with bipolar disorder who are in a euthymia phase, with no differences during a depression phase.
- High quality evidence suggests a small association between poor visual and verbal learning and memory and poor general functioning.
- Moderate quality evidence suggests no changes in list learning over time (~5 years).

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Bo Q, Mao Z, Li X, Wang Z, Wang C, Ma X

Use of the MATRICS consensus cognitive battery (MCCB) to evaluate cognitive deficits in bipolar disorder: A systematic review and meta-analysis

PLoS ONE 2017; 12 (4); doi.org/10.1371/journal.pone.0176212

View review abstract online

Comparison	**Learning in people with bipolar disorder vs. controls.**
Summary of evidence	**High quality evidence (large sample, consistent, precise, direct) suggests medium-sized effects of poorer verbal and visual learning in people with bipolar disorder.**
Verbal and visual learning	
Significant, medium-sized effects of poorer learning in people with bipolar disorder; Verbal learning: 7 studies, N = 487, d = -0.50, 95%CI -0.63 to -0.37, $p < 0.05$, I^2 = 0%, p = 0.76 Visual learning: 7 studies, N = 487, d = -0.57, 95%CI -0.74 to -0.40, $p < 0.05$, I^2 = 39.5%, p = 0.13	
Consistency in results‡	Consistent
Precision in results§	Precise
Directness of results‖	Direct

Bora E, Pantelis C

Meta-analysis of Cognitive Impairment in First-Episode Bipolar Disorder: Comparison With First-Episode Schizophrenia and Healthy Controls

Schizophrenia Bulletin 2015; 41(5): 1095-1104

View review abstract online

Comparison 1	**Learning in people with first-episode bipolar disorder vs. controls.**
Summary of evidence	**Moderate to high quality evidence (large sample, unable to assess consistency, precise, direct) suggests a medium-sized effect of poorer list learning in people with first-episode bipolar disorder.**

List learning	
A significant, medium-sized effect of poorer list learning in people with first-episode bipolar disorder; 5 studies, N = 638, *d* = 0.59, 95%CI 0.40 to 0.78, *p* < 0.001, I^2 not reported. Authors report no evidence of publication bias.	
Consistency in results	Unable to assess; no measure of consistency is reported.
Precision in results	Precise
Directness of results	Direct
Comparison 2	**Learning in people with first-episode bipolar disorder vs. people with first-episode schizophrenia.**
Summary of evidence	**Moderate to high quality evidence (large sample, unable to assess consistency, precise, direct) suggests a medium-sized effect of better list learning in people with first-episode bipolar disorder.**
List learning	
A significant, medium-sized effect of better list learning in people with first-episode bipolar disorder; 4 studies, N = 630, *d* = 0.50, 95%CI 0.14 to 0.87, *p* = 0.007, I^2 not reported. Authors report no evidence of publication bias.	
Consistency in results	Unable to assess; no measure of consistency is reported.
Precision in results	Precise
Directness of results	Direct

Bora E, Ozerdem A

A meta-analysis of neurocognition in youth with familial high risk for bipolar disorder

European Psychiatry 2017; 44: 17-23

View review abstract online

Comparison	**List learning in first-degree relatives aged 10 to 25 years of a person with bipolar disorder vs. controls.**

Summary of evidence	**High quality evidence (large sample, consistent, precise, direct) suggests a small trend effect of poorer list learning in young relatives.**
List learning	
Small trend effect of poorer list learning in relatives; 5 studies, N = 597, *d* = 0.23, 95%CI 0.02 to 0.47, *p* = 0.07, I^2 = 55%, *p* = 0.07	
Consistency	Consistent
Precision	Precise
Directness	Direct

Bora E

Neurocognitive features in clinical subgroups of bipolar disorder: A meta-analysis

Journal of Affective Disorders 2018; 229: 125-34

View review abstract online

Comparison 1	**List learning in people with bipolar I disorder vs. people with bipolar II disorder.**
Summary of evidence	**High quality evidence (large sample, consistent, precise, direct) suggests a small effect of poorer list learning in people with bipolar I disorder.**
List learning	
A significant, small effect of poorer list learning in people with bipolar I disorder; 12 studies, N = 1,355, *d* = 0.28, 95%CI 0.16 to 0.39, *p* < 0.001, I^2 = 0%, *p* = 0.59 There were no significant differences in effect sizes between euthymic and non-euthymic patients.	
Comparison 2	**List learning in people with bipolar disorder and a history of psychotic symptoms vs. people with bipolar disorder and no history of psychotic symptoms.**
Summary of evidence	**High quality evidence (large sample, consistent, precise, direct) suggests a medium-sized effect of poorer list learning in people with bipolar disorder and a history of psychotic symptoms.**

List learning	
A significant, medium-sized effect of poorer list learning in people with bipolar disorder and a history of psychotic symptoms; 6 studies, N = 613, *d* = 0.39, 95%CI 0.22 to 0.55, *p* < 0.001, I^2 = 0%, *p* = 0.83 There were no significant differences in effect sizes between euthymic and non-euthymic or bipolar I and II disorder patients.	
Consistency in results	Consistent
Precision in results	Precise
Directness of results	Direct

Depp CA, Mausbach BT, Harmell AL, Savla GN, Bowie CR, Harvey PD, Patterson TL

Meta-analysis of the association between cognitive abilities and everyday functioning in bipolar disorder

Bipolar Disorders 2012; 14: 217-26

View review abstract online

Comparison	**Associations between learning and functioning in people with bipolar disorder.**
Summary of evidence	**High quality evidence (large sample, consistent, precise, direct) suggests a small association between poor visual and verbal learning and memory and poor general functioning.**
Learning and memory	
Significant, small associations between poor learning and memory and poor general functioning; Verbal learning and memory: 17 studies, N = 895, *r* = 0.23, 95%CI 0.14 to 0.31, *p* < 0.0045, Q*p* = 0.088 Visual learning and memory: 5 studies, N = 381, *r* = 0.26, 95%CI 0.16 to 0.35, *p* < 0.0045, Q*p* = 0.597	
Consistency in results	Consistent
Precision in results	Precise

Directness of results	Direct

Elias LR, Miskowiak KW, Vale AM, Kohler CA, Kjaerstad HL, Stubbs B, Kessing LV, Vieta E, Maes M, Goldstein BI, Carvalho AF

Cognitive Impairment in Euthymic Pediatric Bipolar Disorder: A Systematic Review and Meta-Analysis

Journal of the American Academy of Child & Adolescent Psychiatry 2017; 56: 286-96

View review abstract online

Comparison	**Visual learning and memory in euthymic youth with bipolar disorder vs. controls of similar age (mean 13 years) and IQ (mean 104).**
Summary of evidence	**Moderate quality evidence (small sample, consistent, precise, direct) suggests a large, significant effect of poorer verbal and visual learning and memory in euthymic youth with bipolar disorder.**
Visual learning and memory	
Large, significant effects of poorer visual and verbal learning and memory in euthymic youth with bipolar disorder; Visual learning and memory: 4 studies, N = 135, g = 0.78, 95%CI 0.43 to 1.13, $p < 0.001$, I^2 = 0%, $p > 0.05$ Verbal learning and memory: 7 studies, N = 401, g = 0.76, 95%CI 0.29 to 1.22, p = .001, I^2 = 76%, $p < 0.05$	
Consistency in results	Consistent
Precision in results	Precise
Directness of results	Direct

Samame C, Martino DJ, Strejilevich SA

A quantitative review of neurocognition in euthymic late-life bipolar disorder

Bipolar Disorders 2013; 15: 633-44 View review abstract online	
Comparison	**Learning in older people with bipolar disorder vs. controls matched for age and years of education.**
Summary of evidence	**Moderate to low quality evidence (large sample, inconsistent and imprecise, direct) suggests a medium-sized effect of poorer serial learning in elderly people with bipolar disorder.**
Serial learning	
Medium-sized, significant effect of poorer serial learning in elderly people with bipolar disorder; 3 studies, N = 323, g = 0.76, 95%CI 0.02 to 1.49, p = 0.04, I^2 = 88%, p < 0.001	
Consistency in results	Inconsistent
Precision in results	Imprecise
Directness of results	Direct

Samame C, Martino DJ, Strejilevich SA

Longitudinal course of cognitive deficits in bipolar disorder: a meta-analytic study

Journal of Affective Disorders 2014; 164: 130-8

View review abstract online

Comparison	**Changes in list learning over time in people with bipolar disorder.**
Summary of evidence	**Moderate quality evidence (small samples, consistent, precise, direct) suggests no changes in list learning over time (~5 years).**
List learning	
There were no significant changes over time; 5 studies, N = 166, follow up = 4.83 years, d = 0.00, 95%CI -0.21 to 0.21, p = 0.98, I^2 = 0%, p = 0.43	
Consistency in results	Consistent
Precision in results	Precise

Directness of results	Direct

Samame C, Szmulewicz AG, Valerio MP, Martino DJ, Strejilevich SA

Are major depression and bipolar disorder neuropsychologically distinct? A meta-analysis of comparative studies

European Psychiatry 2017; 39: 17-26

View review abstract online

Comparison	**List learning in people with bipolar disorder vs. people with major depression.**
Summary of evidence	**Moderate quality evidence (small samples, consistent, precise, direct) suggests a medium-sized effect of better list learning in people with major depression compared to people with bipolar disorder who are in a euthymia phase. There were no differences during a depression phase.**
List learning	
A medium-sized significant effect of better list learning in people with major depressive disorder compared to people with bipolar disorder during euthymia only; 3 studies, N = 149, g = 0.65, 95%CI 0.31 to 0.98, $p < 0.001$, I^2 = 0%, p = 0.45 *There were no differences between groups on list learning during depression phases;* 4 studies, N = 151, g = 0.15, 95%CI -0.28 to 0.58, p = 0.50, I^2 = 45%, p = 0.14	
Consistency in results	Consistent
Precision in results	Precise
Directness of results	Direct

Explanation of acronyms

CI = confidence interval, d = Cohen's d and g = Hedges's g standardised mean difference, I^2 = the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance), N = number of participants, p = probability of rejecting a null hypothesis of no differences between groups, Q = test for heterogeneity, r = correlation coefficient

Explanation of technical terms

* Bias has the potential to affect reviews of both RCT and observational studies. Forms of bias include; reporting bias – selective reporting of results; publication bias - trials that are not formally published tend to show less effect than published trials, further if there are statistically significant differences between groups in a trial, these trial results tend to get published before those of trials without significant differences; language bias – only including English language reports; funding bias - source of funding for the primary research with selective reporting of results within primary studies; outcome variable selection bias; database bias - including reports from some databases and not others; citation bias - preferential citation of authors. Trials can also be subject to bias when evaluators are not blind to treatment condition and selection bias of participants if trial samples are small[12].

† Different effect measures are reported by different reviews.

Prevalence refers to how many existing cases there are at a particular point in time. Incidence refers to how many new cases there are per population in a specified time period. Incidence is usually reported as the number of new cases per 100,000 people per year. Alternatively some studies present the number of new cases that have accumulated over several years against a person-years denominator. This denominator is the sum of individual units of time that the persons in the population are at risk of becoming a case. It takes into account the size of the underlying population sample and its age structure over the duration of observation.

Reliability and validity refers to how accurate the instrument is. Sensitivity is the proportion of actual positives that are correctly identified (100% sensitivity = correct identification of all actual positives) and specificity is the proportion of negatives that are correctly identified (100% specificity = not identifying anyone as positive if they are truly not).

Weighted mean difference scores refer to mean differences between treatment and comparison groups after treatment (or occasionally pre to post treatment) and in a randomised trial there is an assumption that both groups are comparable on this measure prior to treatment. Standardised mean differences are divided by the pooled standard deviation (or the standard deviation of one group when groups are homogenous) that allows results from different scales to be combined and compared. Each study's mean difference is then given a weighting depending on the size of the sample and the variability in the data. Less than 0.4 represents a small effect, around 0.5 a medium effect, and over 0.8 represents a large effect[12].

Odds ratio (OR) or relative risk (RR) refers to the probability of a reduction (< 1) or an increase (> 1) in a particular outcome in a treatment group, or a group exposed to a risk factor, relative to the comparison group. For example, a RR of 0.75 translates to a reduction in risk of an outcome of 25% relative to those not receiving the treatment or not exposed to the risk factor. Conversely, a RR of 1.25 translates to an increased risk of 25% relative to those not receiving treatment or not having been exposed to a risk factor. A RR or OR of 1.00 means there is no difference between groups. A medium effect is considered if RR > 2 or < 0.5 and a large effect if RR > 5 or < 0.2[13]. lnOR stands for logarithmic OR where a lnOR of 0 shows no difference between groups. Hazard ratios measure the effect of an explanatory variable on the hazard or risk of an event.

Correlation coefficients (eg, r) indicate the strength of association or relationship

between variables. They can provide an indirect indication of prediction, but do not confirm causality due to possible and often unforseen confounding variables. An r of 0.10 represents a weak association, 0.25 a medium association and 0.40 and over represents a strong association. Unstandardised (*b*) regression coefficients indicate the average change in the dependent variable associated with a 1 unit change in the independent variable, statistically controlling for the other independent variables. Standardised regression coefficients represent the change being in units of standard deviations to allow comparison across different scales.

‡ Inconsistency refers to differing estimates of effect across studies (i.e. heterogeneity or variability in results) that is not explained by subgroup analyses and therefore reduces confidence in the effect estimate. I^2 is the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance) - 0% to 40%: heterogeneity might not be important, 30% to 60%: may represent moderate heterogeneity, 50% to 90%: may represent considerable heterogeneity and over this is considerable heterogeneity. I^2 can be calculated from Q (chi-square) for the test of heterogeneity with the following formula[12];

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%$$

§ Imprecision refers to wide confidence intervals indicating a lack of confidence in the effect estimate. Based on GRADE recommendations, a result for continuous data (standardised mean differences, not weighted mean differences) is considered imprecise if the upper or lower confidence limit crosses an effect size of 0.5 in either direction, and for binary and correlation data, an effect size of 0.25. GRADE also recommends downgrading the evidence when sample size is smaller than 300 (for binary data) and 400 (for continuous data), although for some topics, these criteria should be relaxed[14].

‖ Indirectness of comparison occurs when a comparison of intervention A versus B is not available but A was compared with C and B was compared with C that allows indirect comparisons of the magnitude of effect of A versus B. Indirectness of population, comparator and/or outcome can also occur when the available evidence regarding a particular population, intervention, comparator, or outcome is not available and is therefore inferred from available evidence. These inferred treatment effect sizes are of lower quality than those gained from head-to-head comparisons of A and B.

References

1. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMAGroup (2009): Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *British Medical Journal* 151: 264-9.
2. GRADEWorkingGroup (2004): Grading quality of evidence and strength of recommendations. *British Medical Journal* 328: 1490.
3. Bora E, Pantelis C (2015): Meta-analysis of Cognitive Impairment in First-Episode Bipolar Disorder: Comparison With First-Episode Schizophrenia and Healthy Controls. *Schizophrenia Bulletin* 41: 1095-104.
4. Bora E (2018): Neurocognitive features in clinical subgroups of bipolar disorder: A meta-analysis. *Journal of Affective Disorders* 229: 125-34.
5. Depp CA, Mausbach BT, Harmell AL, Savla GN, Bowie CR, Harvey PD*, et al.* (2012): Meta-analysis of the association between cognitive abilities and everyday functioning in bipolar disorder. *Bipolar Disorders* 14: 217-26.
6. Elias LR, Miskowiak KW, Vale AM, Kohler CA, Kjaerstad HL, Stubbs B*, et al.* (2017): Cognitive Impairment in Euthymic Pediatric Bipolar Disorder: A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child & Adolescent Psychiatry* 56: 286-96.
7. Samame C, Martino DJ, Strejilevich SA (2013): A quantitative review of neurocognition in euthymic late-life bipolar disorder. *Bipolar Disorders* 15: 633-44.
8. 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 [Electronic Resource]* 12 (4): doi.org/10.1371/journal.pone.0176212.
9. Samame C, Martino DJ, Strejilevich SA (2014): Longitudinal course of cognitive deficits in bipolar disorder: a meta-analytic study. *Journal of Affective Disorders* 164: 130-8.
10. Samame C, Szmulewicz AG, Valerio MP, Martino DJ, Strejilevich SA (2017): Are major depression and bipolar disorder neuropsychologically distinct? A meta-analysis of comparative studies. *European Psychiatry* 39: 17-26.
11. Bora E, Ozerdem A (2017): A meta-analysis of neurocognition in youth with familial high risk for bipolar disorder. *European Psychiatry* 44: 17-23.
12. CochraneCollaboration (2008): Cochrane Handbook for Systematic Reviews of Interventions. Accessed 24/06/2011.
13. Rosenthal JA (1996): Qualitative Descriptors of Strength of Association and Effect Size. *Journal of Social Service Research* 21: 37-59.
14. GRADEpro (2008): [Computer program]. Jan Brozek, Andrew Oxman, Holger Schünemann. *Version 32 for Windows*.