

## Cognition and bipolar type

### Introduction

Bipolar disorder is characterised by episodes of depression and mania. A major depressive episode is a period of at least two weeks in which a person has at least five of the following symptoms (including one of the first two): intense sadness or despair; feelings of helplessness, hopelessness or worthlessness; loss of interest in activities once enjoyed; feelings of guilt, restlessness or agitation; sleeping too little or too much; slowed speech or movements; changes in appetite; loss of energy; difficulty concentrating, remembering or making decisions; and/or thoughts of death or suicide.

A manic episode is a period of at least one week when a person is high spirited or irritable in an extreme way most of the day for most days. A manic episode involves changes in normal behaviour such as showing exaggerated self-esteem or grandiosity, less need for sleep, talking more than usual, talking more loudly and quickly, being easily distracted, doing many activities at once, scheduling more events in a day than can be accomplished, embarking on risky behaviour, uncontrollable racing thoughts, and/or quickly changing ideas or topics. These changes in behaviour are significant and clear to friends and family and are severe enough to cause major dysfunction.

The difference between bipolar I disorder and bipolar II disorder is determined by the existence of mania in bipolar I and hypomania in bipolar II. People with bipolar I disorder are also more likely to have psychotic symptoms. As people with psychotic disorders show cognitive deficits, they may be more apparent in people with bipolar I disorder.

### 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<sup>1</sup>. 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 (RCT) 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,



## Cognition and bipolar type

precise and direct with low associated risks (see end of table for an explanation of these terms)<sup>2</sup>. 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 two systematic reviews that met our inclusion criteria<sup>3, 4</sup>.

- Moderate to high quality evidence suggests small effects of greater cognitive impairment in global cognition, verbal memory, processing speed, executive functioning, and language fluency in people with bipolar I disorder compared to bipolar II disorder, with no differences in working or visual memory, attention, inhibition or social cognition.



**Cognition and bipolar type**

*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](#)

<b>Comparison</b>	<b>Cognition in people with bipolar I disorder vs. people with bipolar II disorder.</b>
<b>Summary of evidence</b>	<b>Moderate to high quality evidence (large samples, some inconsistency, precise, direct) suggests small effects of greater cognitive impairment in global cognition, verbal memory, processing speed, and executive functioning in people with bipolar I disorder, with no differences in working or visual memory, attention and social cognition.</b>
<b>Cognitive functioning</b>	
<p><i>Small, significant effects of cognitive impairment in people with bipolar I disorder on;</i></p> <p>Global cognition: 19 studies, N = 1,914, <math>d = 0.17</math>, 95%CI 0.07 to 0.27, <math>p &lt; 0.001</math>, <math>I^2 = 7%</math>, <math>p = 0.38</math></p> <p>Verbal memory: 15 studies, N = 1,459, <math>d = 0.26</math>, 95%CI 0.11 to 0.41, <math>p &lt; 0.001</math>, <math>I^2 = 38%</math>, <math>p = 0.07</math></p> <p>Processing speed: 12 studies, N = 1,325, <math>d = 0.21</math>, 95%CI 0.09 to 0.32, <math>p &lt; 0.001</math>, <math>I^2 = 1%</math>, <math>p = 0.44</math></p> <p>Executive functioning, speed: 14 studies, N = 1,422, <math>d = 0.16</math>, 95%CI 0.00 to 0.32, <math>p = 0.04</math>, <math>I^2 = 48%</math>, <math>p = 0.02</math></p> <p>Executive functioning, accuracy: 11 studies, N = 984, <math>d = 0.17</math>, 95%CI 0.03 to 0.33, <math>p = 0.02</math>, <math>I^2 = 0%</math>, <math>p = 0.53</math></p> <p><i>There were no significant differences on;</i></p> <p>Working memory: 12 studies, N = 1,106, <math>d = 0.02</math>, 95%CI -0.13 to 0.17, <math>p = 0.29</math>, <math>I^2 = 27%</math>, <math>p = 0.18</math></p> <p>Visual memory: 13 studies, N = 967, <math>d = 0.15</math>, 95%CI -0.10 to 0.41, <math>p = 0.23</math>, <math>I^2 = 70%</math>, <math>p = 0.001</math></p> <p>Attention: 7 studies, N = 528, <math>d = 0.17</math>, 95%CI -0.07 to 0.41, <math>p = 0.16</math>, <math>I^2 = 37%</math>, <math>p = 0.15</math></p> <p>Social cognition: 8 studies, N = 464, <math>d = 0.03</math>, 95%CI -0.17 to 0.23, <math>p = 0.76</math>, <math>I^2 = 4%</math>, <math>p = 0.40</math></p> <p>There were no significant moderating effects of symptom severity, number of mood episodes, duration of illness, and antipsychotic use on the effect size for global cognition.</p>	
<b>Consistency in results<sup>†</sup></b>	Consistent, apart from executive functioning - speed, and visual memory.
<b>Precision in results<sup>§</sup></b>	Precise



**Cognition and bipolar type**

<b>Directness of results</b>	Direct
------------------------------	--------

Cotrena C, Damiani Branco L, Ponsoni A, Samame C, Milman Shansis F, Paz Fonseca R

**Executive functions and memory in bipolar disorders I and II: new insights from meta-analytic results**

Acta Psychiatrica Scandinavica 2020; 141: 110-30

[View review abstract online](#)

<b>Comparison</b>	<b>Cognition in people with bipolar I disorder vs. people with bipolar II disorder.</b>
<b>Summary of evidence</b>	<b>Moderate to high quality evidence (large samples, some inconsistency, precise, direct) shows small effects of poorer memory, executive functioning, and language fluency in people with bipolar I disorder, with no differences on inhibition measures.</b>
<b>Memory</b>	
<p><i>Small effects showed people with bipolar I disorder were more impaired than people with bipolar II disorder on;</i></p> <p>Working memory: 8 studies, N = 857, <math>g = 0.22</math>, 95%CI 0.08 to 0.35, <math>p &lt; 0.05</math>, <math>I^2 = 13\%</math>, <math>p = 0.33</math>          Episodic memory: 8 studies, N = 743, <math>g = 0.35</math>, 95%CI 0.15 to 0.55, <math>p &lt; 0.05</math>, <math>I^2 = 60\%</math>, <math>p = 0.01</math>          Immediate verbal episodic memory: 7 studies, N = 520, <math>g = 0.32</math>, 95%CI 0.10 to 0.54, <math>p &lt; 0.05</math>, <math>I^2 = 45\%</math>, <math>p = 0.09</math>          Delayed verbal episodic memory: 6 studies, N = 486, <math>g = 0.30</math>, 95%CI 0.02 to 0.57, <math>p &lt; 0.05</math>, <math>I^2 = 66\%</math>, <math>p = 0.01</math>          Visual episodic memory: 5 studies, N = 536, <math>g = 0.31</math>, 95%CI 0.05 to 0.57, <math>p &lt; 0.05</math>, <math>I^2 = 59\%</math>, <math>p = 0.05</math></p> <p><i>There were no significant differences on;</i></p> <p>Digit span forward: 5 studies, N = 269 365 <math>g = 0.11</math> 95%CI -0.06 to 0.27, <math>I^2 = 0\%</math>, <math>p = 0.56</math></p>	
<b>Executive functioning</b>	
<p><i>Small effects showed people with bipolar I disorder were more impaired than people with bipolar II disorder on;</i></p> <p>Flexibility composite: 10 studies, N = 956, <math>g = 0.22</math>, 95%CI 0.03 to 0.41, <math>p &lt; 0.05</math>, <math>I^2 = 52\%</math>, <math>p = 0.03</math></p>	



**Cognition and bipolar type**

Trail making test B: 7 studies, N = 784, $g = 0.28$ , 95%CI 0.03 to 0.52, $p < 0.05$ , $I^2 = 58%$ , $p = 0.03$	
<b>Language</b>	
<i>Small effects showed people with bipolar I disorder were more impaired than people with bipolar II disorder on;</i>	
Verbal fluency composite: 7 studies, N = 644, $g = 0.18$ , 95%CI 0.04 to 0.31, $p < 0.05$ , $I^2 = 0%$ , $p = 0.64$	
Category fluency: 6 studies, N = 703, $g = 0.22$ , 95%CI 0.06 to 0.37, $p < 0.05$ , $I^2 = 0%$ , $p = 0.66$	
<b>Inhibition/impulsivity</b>	
<i>There were no significant differences on;</i>	
Inhibition composite: 5 studies, N = 442, $g = 0.09$ , 95%CI -0.09 to 0.27, $I^2 = 0%$ , $p = 0.45$	
<b>Consistency in results</b>	Some inconsistency
<b>Precision in results</b>	Precise
<b>Directness of results</b>	Direct

**Explanation of acronyms**

CI = confidence interval,  $d$  or  $g$  = Cohens  $d$  or Hedges'  $g$ , standardised mean differences,  $I^2$  = 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, vs. = versus

## Cognition and bipolar type

### 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<sup>5</sup>.

† 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<sup>5</sup>.

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  $< 0.2$ <sup>6</sup>. InOR stands for logarithmic OR where a InOR 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



## Cognition and bipolar type

between variables. They can provide an indirect indication of prediction, but do not confirm causality due to possible and often unforeseen 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<sup>5</sup>;

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

§ 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<sup>7</sup>.

|| 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.



## Cognition and bipolar type

### References

1. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group (2009): Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *British Medical Journal* 151: 264-9.
2. GRADE Working Group (2004): Grading quality of evidence and strength of recommendations. *British Medical Journal* 328: 1490.
3. Cotrena C, Damiani Branco L, Ponsoni A, Samame 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: 110-30.
4. Bora E (2018): Neurocognitive features in clinical subgroups of bipolar disorder: A meta-analysis. *Journal of Affective Disorders* 229: 125-34.
5. Cochrane Collaboration (2008): Cochrane Handbook for Systematic Reviews of Interventions. Accessed 24/06/2011.
6. Rosenthal JA (1996): Qualitative Descriptors of Strength of Association and Effect Size. *Journal of Social Service Research* 21: 37-59.
7. GRADEpro (2008): [Computer program]. Jan Brozek, Andrew Oxman, Holger Schünemann. *Version 3.2 for Windows*