TECHNICAL COMMENTARY

NeuRA Discover. Conquer. Cure.
BIPOLAR DISORDERS LIBRARY

Reasoning and problem solving

Introduction

Reasoning refers to the ability to logically gather information to solve problems and form conclusions. Reasoning bias may affect problem solving skills and is measured in three ways: 'jumping to conclusions' (JTC); 'belief inflexibility'; and an 'externalising attribution style'. JTC can be measured with the Bead task that presents participants with two jars containing different ratios of coloured beads (eg. 80 red: 20 blue). Beads are drawn from one of the jars, and based on the string of coloured beads drawn, participants must guess which jar they were drawn from. Within the JTC task, "draws to decision" refers to the number of beads required to decide which jar they were drawn from. Extreme JTC responding refers to when a decision is made after little information is gathered. The "draws to certainty" condition is when participants are asked about their certainty regarding which jar beads are being drawn from. "Response to disconfirmatory evidence" refers to the change in certainty after a single bead contradicts their response. "Response to reversal" is when a participant makes a decision based on the initial evidence, then reverses their decision based on later evidence. Belief inflexibility is an inability to change a belief when presented with contradictory evidence, and can be measured by the Bias Against Disconfirmatory Evidence (BADE) task. Attribution bias refers to when available evidence is incorrectly used to attribute an event to internal or external causes and is measured by the Pragmatic Inference Task or Attribution questionnaire where participants are asked to explain events. Reasoning and problem solving may also be measured using Mazes or the Matrix Reasoning where participants select the missing design in a patterned sequence.

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

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

objective summary of the available evidence, although the conclusions are solely the opinion of staff of NeuRA (Neuroscience Research Australia).

Results

We found four systematic reviews that met our inclusion criteria[3-6].

- High quality evidence suggests a small effect of poorer reasoning and problem solving in people with bipolar disorder (including first-episode) compared to controls.
- Moderate quality evidence shows no significant differences in reasoning ability between people with first-episode bipolar disorder and people with first-episode schizophrenia.
- Moderate to high quality evidence suggests a small association between poor reasoning/problem solving and poor general functioning.
- Moderate quality evidence suggests no differences in reasoning/problem solving between euthymic youth with bipolar disorder and age and IQ-matched controls.

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	**Reasoning and problem solving in people with bipolar disorder vs. controls.**
Summary of evidence	**High quality evidence (large sample, consistent, precise, direct) suggests a small effect of poorer reasoning and problem solving in people with bipolar disorder.**
Reasoning and problem solving	
A significant, small effect of poorer reasoning and problem solving in people with bipolar disorder; 7 studies, N = 487, d = -0.30, 95%CI -0.43 to -0.17, $p < 0.05$, I^2 = 6.3%, p = 0.379	
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	**Reasoning and problem solving in people with first-episode bipolar disorder vs. controls.**
Summary of evidence	**Moderate to high quality evidence (large sample, inconsistent, precise, direct) shows a small effect of poorer reasoning in people with first-episode bipolar disorder.**

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

Reasoning	
A significant, small effect of poorer reasoning in people with first-episode bipolar disorder; 7 studies, N = 1,053, *d* = 0.31, 95%CI 0.05 to 0.56, *p* = 0.02, I^2 = 62%, *p* = 0.01 Authors report no evidence of publication bias.	
Consistency in results	Inconsistent
Precision in results	Precise
Directness of results	Direct
Comparison 2	**Reasoning and problem solving in people with first-episode bipolar disorder vs. people with first-episode schizophrenia.**
Summary of evidence	**Moderate quality evidence (small to medium-sized sample, consistent, precise, direct) shows no significant differences in reasoning ability.**
Reasoning	
No significant differences in reasoning; 2 studies, N = 218, *d* = 0.23, 95%CI −0.09 to 0.56, *p* = 0.16, I^2 = 26.3%, *p* = 0.24 Authors report no publication bias. No differences were found for males vs. females or younger vs. older 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 reasoning and problem solving and

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

	functioning in people with bipolar disorder.
Summary of evidence	**Moderate to high quality evidence (medium-sized sample, consistent, precise, direct) suggests a small association between poor reasoning/problem solving and poor general functioning.**
Reasoning and problem solving	
Significant, small association between poor reasoning and problem solving and poor general functioning; 10 studies, N > 247, r = 0.23, 95%CI 0.14 to 0.32, p < 0.0045, Qp = 0.686	
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	**Reasoning and problem solving in euthymic youth with bipolar disorder vs. controls of similar age (mean 13 years) and IQ (mean 104).**
Summary of evidence	**Moderate quality evidence (large sample, inconsistent, imprecise, direct) suggests no differences in reasoning/problem solving.**
Reasoning and problem solving	
No significant differences were found; 12 studies, N = 525, d = 0.43, 95%CI -0.12 to 0.97, p = 0.125, I^2 = 88%, p < 0.05	
Consistency in results	Inconsistent
Precision in results	Imprecise

Directness of results	Direct

Explanation of acronyms

CI = confidence interval, *d* = Cohen's d 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

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

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[7].

---

† 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[7].

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[8]. 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

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

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[7];

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

§ 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[9].

‖ 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.

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

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. 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.
5. 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.
6. 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.
7. CochraneCollaboration (2008): Cochrane Handbook for Systematic Reviews of Interventions. Accessed 24/06/2011.
8. Rosenthal JA (1996): Qualitative Descriptors of Strength of Association and Effect Size. *Journal of Social Service Research* 21: 37-59.
9. GRADEpro (2008): [Computer program]. Jan Brozek, Andrew Oxman, Holger Schünemann. *Version 32 for Windows*.

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