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

Liana R. Elias, PhD, Kamilla W. Miskowiak, PhD, Antônio M.O. Vale, MSc, Cristiano A. Köhler, PhD, Hanne L. Kjærstad, MSc, Brendon Stubbs, PhD, Lars V. Kessing, MD, PhD, Eduard Vieta, MD, PhD, Michael Maes, MD, PhD, Benjamin I. Goldstein, MD, PhD, André F. Carvalho, MD, PhD

PII: S0890-8567(17)30021-7
DOI: 10.1016/j.jaac.2017.01.008
Reference: JAAC 1697

To appear in: Journal of the American Academy of Child & Adolescent Psychiatry

Received Date: 26 August 2016
Revised Date: 21 December 2016
Accepted Date: 11 January 2017


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Cognitive Impairment in Euthymic Pediatric Bipolar Disorder: A Systematic Review and Meta-Analysis

RH: Neurocognition in Euthymic BD
Liana R. Elias, PhD; Kamilla W. Miskowiak, PhD; Antônio M.O. Vale, MSc; Cristiano A. Köhler, PhD; Hanne L. Kjærstad, MSc; Brendon Stubbs, PhD; Lars V. Kessing, MD, PhD; Eduard Vieta, MD, PhD; Michael Maes, MD, PhD; Benjamin I. Goldstein, MD, PhD; André F. Carvalho, MD, PhD
Clinical guidance is available at the end of this article.
Supplemental material cited in this article is available online.
Accepted January 25, 2017

Drs. Elias, Köhler, Carvalho, and Mr. Vale are with Translational Psychiatry Research Group Federal University of Ceará, Fortaleza, CE, Brazil. Drs. Miskowiak, Kessing, and Ms. Kjærstad are with Psychiatric Centre Copenhagen, Copenhagen University Hospital, Copenhagen, Denmark. Dr. Stubbs is with Health Service and Population Research Department, Institute of Psychiatry, King’s College London, UK, and South London and Maudsley National Health Service Foundation Trust, Denmark Hill, London. Dr. Vieta is with Bipolar Disorder Program, Institute of Neuroscience, Hospital Clinic, University of Barcelona, Institut d’investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Centro para la Investigación Biomédica en Red de Salud Mental (CIBERSAM), Barcelona, Spain. Dr. Maes is with Chulalongkorn University, Bangkok, Thailand, and IMPACT Strategic Research Center, Barwon Health, Deakin University, Geelong, VIC, Australia. Dr. Goldstein is with Centre for Youth Bipolar Disorder, Sunnybrook Health Sciences Centre, Toronto, Canada.

Dr. Elias is supported by a doctoral research fellowship from the Fundação Cearense de Apoio à Pesquisa (FUNCAP; Brazil). Dr. Köhler is supported by a postdoctoral research fellowship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES; Brazil). Dr. Carvalho is supported by a research fellowship award from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq; Brazil). Dr. Miskowiak received support from the Lundbeck Foundation and the Weimann Stipend for half of her salary for her to do full-time research from 2012-2018.

Drs. Elias and Dr. Miskowiak contributed equally as first authors.

The authors thank Ana Kleinman, MD, PhD from the University of São Paulo (Brazil), Anne Helseth Udal, MD, PhD from the Sørlandet Hospital (Norway), Judah David
Weathers, MD, PhD from the Department of Psychiatry at Yale School of Medicine, Ellen Leibenluft, MD from the National Institute of Mental Health (US), and Rodrigo Grassi-Oliveira, MD, PhD from the Pontifical Catholic University of Rio Grande do Sul (Brazil) for kindly providing additional data for this review.

Disclosure: Dr. Miskowiak has received consultancy fees from Lundbeck and Allergan. Dr. Kessing has served as a consultant for Lundbeck, Astra Zeneca, and Sunovion. Dr. Vieta has received research grants, CME-related honoraria, or consulting fees from: ABBiotics, Allergan, AstraZeneca, Ferrer, Forest Research Institute, Gedeon Richter, GlaxoSmith-Kline, Janssen, Lundbeck, Otsuka, Pfizer, Sanofi-Aventis, Sunovion, and Takeda, as well as from CIBERSAM, Grups Consolidats de Recerca 2014 (SGR 398), Seventh European Framework Programme (ENBREC), and the Stanley Medical Research Institute. Dr. Goldstein has received grant or research support from Brain Canada, the Canadian Institutes of Health Research, the Brain and Behavior Research Foundation (NARSAD), the National Institute of Mental Health, the Ontario Mental Health Foundation, and the Ontario Ministry of Research and Innovation. Drs. Elias, Vale, Köhler, Kjærstad, Stubbs, Maes, and Carvalho report no biomedical financial interests or potential conflicts of interest.

Correspondence to André F. Carvalho, MD, PhD, Department of Clinical Medicine, Faculty of Medicine, Federal University of Ceará, Rua Prof. Costa Mendes, 1608, 4º andar, 60430-040, Fortaleza, CE, Brazil; email: andrefc7@terra.com.br.
ABSTRACT

Objective: To perform a systematic review and meta-analysis of studies investigating neurocognition in euthymic youths with bipolar disorder (BD) compared to healthy controls (HCs).

Method: A systematic literature search was conducted in the PubMed/MEDLINE, PsycINFO, and EMBASE databases from inception up until March 23rd, 2016 for original peer-reviewed articles that investigated neurocognition in euthymic youths with BD compared to HCs. Effect sizes (ES) for individual tests were extracted. In addition, results were grouped according to cognitive domain. This review complied with the PRISMA statement.

Results: 24 studies met inclusion criteria (N=1,146; 510 with BD). Overall, euthymic youths with BD were significantly impaired in verbal learning, verbal memory, working memory, visual learning, and visual memory with moderate to large ESs (Hedge’s g between 0.76 to 0.99); significant impairments were not observed for attention/vigilance, reasoning and problem solving, and/or processing speed. Heterogeneity was moderate to large ($I^2 \geq 50\%$) for most ES estimates. Differences in the definition of euthymia across studies explained the heterogeneity in the ES estimate for verbal learning and memory. We also found evidence for other potential sources of heterogeneity in several ES estimates including co-occurring attention-deficit/hyperactivity disorder (ADHD) and anxiety disorders, and the use of medications. In addition, the use of different neuropsychological tests appeared to contribute to heterogeneity of some estimates (e.g. attention/vigilance domain).

Conclusion: Euthymic youths with BD exhibit significant cognitive dysfunction encompassing verbal learning and memory, working memory, and/or visual learning and memory domains. These data indicate that for a subset of individuals with BD, neurodevelopmental factors may contribute to cognitive dysfunction.
**Key words:** bipolar disorder, cognition, meta-analysis, psychiatry, memory

**INTRODUCTION**

Evidence that bipolar disorder (BD) in adults is associated with broad neurocognitive impairment across several domains has accumulated in the literature in the past two decades.\(^1\)\(^,\)\(^2\) These findings have been synthesized in several meta-analyses\(^3\)\^-\(^5\) that confirmed that euthymic adults with BD have most pronounced deficits (of moderate to large effect sizes) in attention, verbal memory, and executive function. Furthermore, a large individual patient meta-analysis confirmed that euthymic adults with BD exhibit cognitive impairment encompassing several domains.\(^6\) Nevertheless, this effort found smaller effect sizes in comparison to previous meta-analyses, perhaps due to a more accurate adjustment for potential confounders and the inclusion of unpublished datasets.\(^6\)

BD has been increasingly diagnosed in pediatric age groups, with an estimated pooled prevalence rate of 1.8% in general population surveys.\(^7\) In addition, a significant proportion of adults with BD have onset of illness before the age of 18.\(^8\)\^-\(^9\) Patients with early onset of illness may have worse outcomes including but not limited to fewer days of euthymia and greater functional impairment, which may hamper academic achievements.\(^9\)\^-\(^12\)

Cognitive impairment is thought to be a main mediator of disability and functional impairment in BD.\(^2\)\(^,\)\(^13\) Evidence indicates that cognitive impairment may be evident in first-episode BD.\(^14\) A previous systematic review\(^15\) and two meta-analyses\(^16\)\^-\(^17\) suggested that cognitive impairment in pediatric BD is mainly present in verbal learning, processing speed, and executive function domain. However, these analyses included individuals with BD during different mood states, and the largest available meta-analysis included 12 unique studies (N=374).\(^17\) Since the publication of those previous systematic reviews and meta-analyses, several new original studies provided data on cognitive function in euthymic youths with BD.\(^18\)\^-\(^21\) However, different clinical characteristics of samples may lead to heterogeneity
across studies. For example, evidence suggests that the use of medications\textsuperscript{21} and co-occurring mental disorders (notably attention-deficit/hyperactivity disorder [ADHD])\textsuperscript{22} may impact cognitive function in youths with BD.

Therefore, a systematic review and meta-analysis of studies that had assessed cognitive impairment in euthymic youths with BD compared to healthy controls was performed. We expected a high degree of heterogeneity across studies in part due to differences in sample characteristics (for example, duration of illness, comorbid ADHD, and medication status) as well as differences in neuropsychological measures\textsuperscript{23}. Thus, we also aimed to explore potential sources of heterogeneity.

**METHOD**

This study comprised a meta-analysis of studies comparing neurocognitive function in euthymic youths with BD to healthy controls. We complied with the Preferred Reported Items for Systematic Reviews and Meta-analysis (PRISMA) statement, and followed an a priori defined yet unpublished protocol\textsuperscript{24}. The literature search, title/abstract screening, final decision on eligibility after full-text review, and data extraction were independently performed by at least two investigators (L.R.O., A.M.O.V., and H.L.K.). Disagreements were resolved through consensus. If a consensus could not be achieved, the decision was made independently by a third investigator (C.A.K.).

**Search Strategy**

A systematic search was conducted in the EMBASE, PubMed/MEDLINE, and PsycInfo databases from inception up until March 23\textsuperscript{rd}, 2016. The detailed search strings used in this review are presented in the supplementary online material that accompanies the online version of this article. This search strategy was augmented through tracking the citation of included articles in Google Scholar\textsuperscript{25}.

**Study Selection**
We included original peer-reviewed articles published in English, Danish, Spanish, Portuguese, French, or German. Eligible studies had to include youth (i.e., aged ≤ 18 years old) participants meeting either DSM26 or International Classification of Diseases (ICD)27 criteria for BD, and a comparison sample of healthy controls (HCs). In addition, the diagnosis of BD had to be established through a validated structured/semi-structured diagnostic interview. For this systematic review, we considered studies that had measured cognitive domains according to standard tests and/or batteries (vide infra). Neurocognitive function has been heuristically subdivided into “hot” (i.e., emotion-laden) and “cold” (i.e., emotion-independent) cognition.28, 29 For the purposes of this systematic review, we had solely included studies measuring “cold” cognition because this dimension has been more consistently related to clinical outcomes (e.g., functional status).15 We applied the following exclusion criteria: (1) studies that did not diagnose BD by means of a structured diagnostic interview; (2) meeting abstracts; (3) studies published in other languages (i.e., not specified above); (4) studies that included participants from other age groups (unless data were reported separately for youths); (5) investigations that included participants with BD in an acute mood episode or otherwise clinically unstable; (6) studies that did not include an HC group.

Data Extraction

We extracted the means ± standard deviation (SD) of each neuropsychological test score or the appropriate reported effect size estimate. In addition to meta-analysis of individual tests, we also grouped tests according to cognitive domains based on the National Institute of Mental Health-MATRICS (Measurement and Treatment Research to Improve Outcome in Schizophrenia) with slight modifications due to available data from tests not included in this initiative (Table S1, available online).30, 31 We also extracted the following data whenever available: (1) First author; (2) Publication Year; (3) sample size; (4) gender.
distribution (% Females); (5) mean illness duration (years); (6) definition of euthymia considered in each study; (7) the medications in use (categorized as lithium, anticonvulsants, and antipsychotics); and (8) overall intelligence (mean IQ scores). Evidence points to a high prevalence of comorbid psychiatric conditions (up to 90%) among youths with BD.\textsuperscript{33-35} In addition, some reports indicate that a significant proportion of youths with BD may have comorbid ADHD in clinical settings.\textsuperscript{36,37} Therefore, we extracted data on the frequency of comorbid mental disorders in each included study; we considered cognitive data for participants with BD and no comorbid mental disorders whenever a study provided these data.

**Statistical Analysis**

Because studies used different measurement methods, standardized mean difference estimates of difference in test scores were used as effect size (ES) estimates utilizing Hedge’s $g$, which provides an unbiased ES adjusted for small sample sizes.\textsuperscript{38} The 95% CI was also computed.\textsuperscript{39} An ES of 0.2 was considered low, 0.5 moderate, and 0.8 large.\textsuperscript{40} The ES of each cognitive domain represents the average of the ES estimates derived from each of its pertinent neuropsychological measures.\textsuperscript{41} We also calculated a global cognition measure by averaging the ES from each cognitive domain, when three or more domains were assessed in the same study.\textsuperscript{41} This composite measure aimed to provide a broad estimate of overall cognitive function. At least three independent datasets had to be available to estimate a summary ES (for both individual tests and cognitive domains).

Studies with statistically non-significant (i.e. negative) results are less likely than studies with positive results of being published.\textsuperscript{42,43} To account for significant publication bias, we inspected a funnel plot graph for asymmetry and calculated the Egger’s test.\textsuperscript{44}

We assessed the heterogeneity across studies using the Cochran $Q$ test, a weighted sum of the squares of the deviations of individual study ES estimates from the overall
estimate. In addition, heterogeneity across studies was quantified with the $I^2$ statistic, which in brief indicates the percentage of total variation across several studies due to heterogeneity, and it is considered moderate when between 50-75%, and high when greater than 75%.\textsuperscript{45} We anticipated a high degree of heterogeneity. Therefore, we pooled ES from several studies according to the inverse variance method of accounting for random effects. Random-effects modelling assumes a genuine diversity across studies and incorporates a between-study variance into the calculations.\textsuperscript{39}

We explored potential sources of heterogeneity across studies in each cognitive domain with either subgroup or random-effects meta-regression analyses. We considered the following variables in univariate meta-regression analyses: sample size, mean age of BD group, mean age of the HC group, differences in mean age (BD group minus HC group), % of females in the BD group, % of females in the HC group, difference in % of females (BD group minus HC group), mean IQ of the BD group, mean IQ of the HC group, difference in mean IQ (BD group minus HC group), % of participants on lithium, % of participants on anticonvulsants, % of participants on antipsychotics. Studies were weighted in such a way that investigations with more precise parameters (indicated by sample size and 95% CI) had more influence in meta-regression analyses.\textsuperscript{46} A covariate was investigated in meta-regression analyses when at least five independent datasets provided data on the potential moderator. We performed a subgroup analysis based on the criteria of euthymia employed in each study. For the purposes of these analyses, we classified each study based on a stringent definition of euthymia (i.e., euthymia as pre-defined based on scores in standard symptom rating instruments). A non-stringent definition of euthymia was considered when at least one of the following were met: (1) mean scores of the Young Mania Rating Scale (YMRS)\textsuperscript{47} < 12 and the Children’s Depression Rating Scale-Revised (CDRS-R)\textsuperscript{48} ≤ 28 or (2) euthymia was established by a clinical assessment. We electronically contacted the corresponding author of
manuscripts in at least two separate occasions if there was uncertainty regarding the definition of euthymia. In addition, we performed subgroup analyses on the following variables: (1) presence of comorbid ADHD in the BD group (Yes/No); (2) presence of comorbid anxiety disorders in the BD group (Yes/No); (3) history of psychotic symptoms; (4) on lithium use; (5) on antipsychotic use; and (6) on anticonvulsant use. We estimated the power of each summary ES estimate across a hypothetical range of ‘true’ ESs with a method developed by Hedges and Pigott.49

All analyses were conducted with the metan package in Stata MP software version 14.0 (Stata-Corp, College Station, TX, USA). Statistical significance was considered at an alpha level of 0.05.

RESULTS

Following the removal of duplicates, the title/abstracts of 1,909 unique references were screened for eligibility. A total of 1,675 references were excluded, while 234 full-text were retrieved and screened for eligibility. Finally, 24 original studies met inclusion criteria.18-22,50-68 Figure 1 provides the PRISMA flowchart for study selection.

Characteristics of Included Studies

This systematic review and meta-analysis included 24 studies that provided data from 1,146 participants (510 youths with BD and 636 HCs). Across the 24 included studies, there were no differences in age (BD group: 13.62 ± 3.76; HC group: 13.29 ± 3.22, P = .12) and IQ scores (BD group: 103.58 ± 15.76; HC group: 104.64 ± 24.32, P = .44) between groups, while 12 studies followed an age- and gender-matched design. Table S2 (available online) provides a description of included studies.

Global Cognition

Seven studies provided evidence that youths with BD are significantly impaired in global cognition (Table 1). There was no evidence of publication bias. Heterogeneity was
large ($I^2 = 85\%$). Figure 2A provides the forest plot of this meta-analysis. Mean age of the BD group appeared to moderate this ES estimate, with older participants being more cognitively impaired. The frequency of current use of lithium and antipsychotics also moderated this outcome, where use of these medications appeared to impair this domain (Table S3, available online). Subgroup analyses found that global cognition was impaired in studies that included youths with BD and comorbid anxiety disorders (with low heterogeneity) but not in those that included participants without comorbid anxiety disorders. In addition, deficits in this domain were significant in studies that included youths with BD and co-morbid ADHD (with low heterogeneity) but not in those that included youths with BD without comorbid ADHD. Lastly, impairment in this domain was observed only in studies that considered the stringent criteria for euthymia albeit with high heterogeneity (Table S4, available online).

<Please insert Table 1 here>

**Attention and Vigilance**

Overall, attention and vigilance were not significantly impaired in youths with BD compared to HCs (Table 1; Figure 3). However, a moderate degree of heterogeneity was observed for this ES estimate. Meta-regression analyses suggested that the mean age and the percentage of female participants in the HC moderated this outcome. Furthermore, the frequency of current use of anticonvulsants and antipsychotics emerged as moderators of this outcome, with evidence suggesting that the use of these drugs could impair this cognitive domain (Table S3, available online). Subgroup analyses suggest that this cognitive domain could be impaired in youths with BD in studies that considered the stringent criteria for euthymia (with low heterogeneity) but not in studies that considered other definitions (Table S4, available online). Furthermore, impairment in this domain was observed in studies that included youths with BD and comorbid anxiety disorders, but not studies that included only
Participants with BD were slightly impaired in the hit reaction time and the omission scores of the Continuous Performance Test-II (CPT-II). These results suggest that youths with BD may have difficulties with sustained attention. Interestingly, heterogeneity for this ES estimates were low. In addition, results of the Go/Nogo tasks seem to confirm that youths with BD have deficient impulse control (Table 1).

**Reasoning and Problem Solving**

Youths with BD were not significantly impaired in reasoning and problem solving (Table 1; Figure 4). However, heterogeneity for this ES estimate was high. Meta-regression analyses indicated that the frequency of use of antipsychotics and anticonvulsants moderated this ES, with data suggesting that the use of these drugs could impair this domain (Table S3, available online).

**Verbal Learning and Memory**

Youths with BD were significantly impaired in verbal learning and memory compared to HCs, with a moderate ES (Figure 2B). Heterogeneity was large (Table 1). In studies that considered a stringent definition for euthymia, heterogeneity was low and a significant impairment was observed, whereas no significant effect was verified in studies that used other definitions of euthymia. Furthermore, impairment in this domain was observed in studies that included youths with BD and co-occurring ADHD (with low heterogeneity) but not in studies that included solely youths with BD without this comorbidity (Table S4, available online).

**Working Memory**

Our results indicate that pediatric BD is associated with large impairments in working memory (Table 1; Figure 2C). Heterogeneity was moderate ($I^2=62\%$). However, it is noteworthy that results obtained with the Wechsler Intelligence Scale for Children (WISC)
digit span forward test confirmed this finding, with low heterogeneity (Table 1). Heterogeneity was large in studies that included youths with BD and comorbid anxiety disorders, but low in studies that included youths with BD without this comorbidity (Table S4, available online).

**Speed of Processing**

Just three studies provided extractable data for the speed of processing domain.\(^{22,51,65}\) Evidence suggests that youths with BD are not significantly impaired in this domain (Table 1; Figure 5).

**Visual Learning and Memory**

Youths with BD were significantly impaired in visual learning and memory (Table 1; Figure 2D) with a moderate to large ES (Hedge’s \(g = 0.78\)). In addition, heterogeneity was low.

**Verbal Fluency**

A single small study (N=40) investigated this cognitive domain by means of the Benton Controlled Oral Association Test, phonemic Fluency FAS (COWAT).\(^{21,70}\) No significant difference between euthymic adolescents with BD and age- and gender-matched controls was observed.

**Power Analysis**

A post hoc power analysis indicate that the meta-analysis of the attention and vigilance domain had adequate power to reject the null hypothesis considering a medium ‘true’ effect size, whereas the summary ES estimates for reasoning and problem solving and speed of processing had a low power to reject the null hypothesis considering a medium ‘true’ effect size (Table S5, available online).

**DISCUSSION**

This meta-analysis provides evidence that euthymic pediatric patients with BD have
significant cognitive impairment of moderate-to-large ES when compared to HCs in the global cognition score, as well as in verbal and visual learning and memory and working memory domains. These results are consistent with previous meta-analyses conducted in euthymic adults with BD that provided evidence that these domains are impaired when compared to HCs.\textsuperscript{3,5,71} However, available evidence indicates that other cognitive domains are similarly impaired in euthymic adults with BD, including attention/vigilance, reasoning and problem solving, and processing speed.\textsuperscript{5} These results suggest that cognitive dysfunction may be broader in adults compared to youths with BD.

Two recent meta-analyses indicate that global cognitive deficits are already evident in first-episode BD.\textsuperscript{14, 72} However, these meta-analyses included only studies of adult samples. Our results provide evidence that cognitive dysfunction encompassing several cognitive domains could already be evident in euthymic youths with BD.

Some limitations of our work deserve discussion. First, heterogeneity for several ES estimates was high. We explored potential sources of heterogeneity across studies with standard meta-regression and subgroup analyses. Results from meta-regressions should be cautiously interpreted due to the possibility of type I error.\textsuperscript{73} We found evidence that the use of medications contributed to the heterogeneity of most ES estimates. In addition, meta-regression analyses suggested that the frequency of use of lithium and antipsychotics may have a detrimental effect upon global cognition. Few studies have investigated the cognitive effects of those medications in youths with BD. A previous observational study found that mood stabilizers may adversely impact certain neurocognitive domains, while atypical antipsychotics had no significant effect on neurocognition.\textsuperscript{74} Clearly our findings should be interpreted cautiously due to the inclusion of few studies in those analyses, and future prospective studies are needed to investigate the impact of mood stabilizing medications on cognition in pediatric samples with BD. In addition, it is possible that the effects observed in
our exploratory analyses could be confounded by other clinical variable (e.g., youths with more severe BD could more likely require the use of mood stabilizing medications). We could not control for these covariates (e.g., length of illness, number of affective episodes) due to the fact that included studies did not consistently provide data. It is worthy to note that the definition of euthymia has varied across studies, which might have influenced some findings of our analyses. For example, most of the heterogeneity in the ES estimate of the verbal learning and memory domain appeared to be explained by the criterion of euthymia adopted by each study, with low heterogeneity for studies that used a more stringent definition of euthymia. Furthermore, our analyses suggest that co-morbid ADHD or anxiety disorders contributed to the heterogeneity of several estimates, and we found some indications that these co-occurring disorders may negatively impact some neurocognitive domains (e.g. global cognition and attention/vigilance). These suggestive findings add to emerging evidence that indicates that these comorbidities may have a deleterious impact on the course of pediatric BD.

Lastly, neuropsychological measures tapping into the cognitive domains varied across studies. This aspect may have contributed to heterogeneity of some estimates. For instance, whilst no significant differences emerged in the attention/vigilance domain, individual test meta-analyses (in the CPT-II and Go/Nogo tasks) suggested that youths with BD do display impaired attention and impulse control regulation.

Two previous quantitative analyses have included youths with BD across the spectrum of mood states. Our meta-analysis was restricted to the inclusion of euthymic pediatric participants with BD, thus allowing a more precise estimation of trait-related cognitive deficits. Furthermore, this updated meta-analysis included twice the number of studies compared to the largest previous meta-analysis. This previous meta-analysis found broader cognitive impairments (with medium ES estimates) in youths with BD encompassing general cognitive ability, attention, executive control, working memory, visuospatial skills,
verbal fluency, verbal learning and memory, and visual memory domains. However, the inclusion of participants in different mood states might have confounded their results as suggested by an exploratory moderator analysis performed by the authors. Thus, an important strength of our work rests in the provision of the best possible synthesis of available evidence.

Cognitive impairment has emerged as a consistent target for drug and psychological treatments for BD. Our results provide a rationale for testing the impact of those interventions in youths with BD. In addition, the impact of cognitive dysfunction in functioning among youths with BD remains largely unknown although a controlled study indicates that executive deficits may impair academic achievement in youths with BD.

This meta-analysis has several research implications. First, our results indicate that a better standardization of neuropsychological assessments across studies investigating neurocognitive dysfunction in youths with BD is necessary. Our results may provide useful indications for this purpose. For example, our individual test meta-analysis suggests that the WISC-digit span forward might provide a sensitive and consistent (with low heterogeneity) measure of working memory in youths with BD. Second, future studies should provide a better description and control for potential confounding variables. Third, more studies are necessary to investigate cognitive dysfunction encompassing the processing speed and verbal fluency domains in euthymic youths with BD. Fourth, available evidence does not consistently indicate that adult patients with early-onset BD could be more cognitively impaired after adjustment for potential confounders. Therefore, prospective studies are warranted to precisely determine the course of cognitive deficits in youths with BD. Finally, future studies should control for the presence of co-occurring ADHD as well as compare cognitive dysfunction in youths with BD to youths with ADHD without comorbid mood disorders.
Clinical guidance

- This meta-analysis indicates that euthymic youths with bipolar disorder exhibit impairments in global cognition, verbal and visual learning and memory, and working memory compared to healthy controls with moderate to large effect sizes.

- Neuropsychological assessments have varied across studies. In addition, the definition of euthymia was not consistent in included investigations. The impact of some illness variables (e.g. illness duration and number of previous affective episodes) on cognitive dysfunction in pediatric BD remains incompletely elucidated.

- Our findings indicate that cognitive dysfunction may be a viable therapeutic target for pharmacological and psychological intervention in euthymic youths with BD.

FIGURE LEGENDS

FIGURE 1. Preferred Reported Items for Systematic Reviews and Meta-analysis (PRISMA) flowchart of included studies. Note: MATRICs = Measurement and Treatment Research to Improve Outcome in Schizophrenia.

FIGURE 2. Forest plots for effect size estimates (and 95% CI) for (A) global cognition; (B) verbal learning and memory; (C) working memory; and (D) visual learning and memory in euthymic youths with bipolar disorder (BD) compared to healthy controls (HCs).

FIGURE 3. Forest plots for effect size estimates (and 95% CIs) for the attention and vigilance cognitive domain. Note: BD = bipolar disorder.

FIGURE 4. Forest plots for effect size estimates (and 95% CIs) for the reasoning and problem solving cognitive domain. Note: BD = bipolar disorder.

FIGURE 5. Forest plots for effect size estimates (and 95% CIs) for the speed of processing cognitive domain. Note: BD = bipolar disorder.
<table>
<thead>
<tr>
<th>Domain/Test</th>
<th>N studies</th>
<th>N cases</th>
<th>N healthy controls</th>
<th>Hedges's g</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
<th>Overall P-value</th>
<th>Q</th>
<th>I² (%)</th>
<th>Small-study effects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Global cognition</strong></td>
<td>7</td>
<td>132</td>
<td>210</td>
<td>0.78</td>
<td>0.16</td>
<td>1.41</td>
<td><strong>.014</strong></td>
<td><strong>38.63</strong></td>
<td>85</td>
<td>0.349</td>
</tr>
<tr>
<td><strong>Attention/vigilance</strong></td>
<td>12</td>
<td>254</td>
<td>349</td>
<td>0.18</td>
<td>-0.06</td>
<td>0.43</td>
<td><strong>.144</strong></td>
<td><strong>22.11</strong></td>
<td>50</td>
<td>0.734</td>
</tr>
<tr>
<td>Continuous Performance Test-II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comissions</td>
<td>4</td>
<td>68</td>
<td>196</td>
<td>-0.14</td>
<td>-0.51</td>
<td>0.23</td>
<td>.462</td>
<td>5.15</td>
<td>42</td>
<td>0.644</td>
</tr>
<tr>
<td>Hit reaction time</td>
<td>4</td>
<td>67</td>
<td>148</td>
<td>0.48</td>
<td>0.09</td>
<td>0.86</td>
<td><strong>.017</strong></td>
<td>4.74</td>
<td>37</td>
<td>0.660</td>
</tr>
<tr>
<td>Omissions</td>
<td>5</td>
<td>92</td>
<td>173</td>
<td>0.41</td>
<td>0.15</td>
<td>0.67</td>
<td><strong>.002</strong></td>
<td>3.68</td>
<td>0</td>
<td>0.528</td>
</tr>
<tr>
<td>Go-Nogo tasks</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy go (%)</td>
<td>3</td>
<td>76</td>
<td>63</td>
<td>0.03</td>
<td>-0.31</td>
<td>0.38</td>
<td>.854</td>
<td>2.17</td>
<td>8</td>
<td>0.223</td>
</tr>
<tr>
<td>Go reaction time (ms)</td>
<td>3</td>
<td>76</td>
<td>63</td>
<td>-0.36</td>
<td>-0.69</td>
<td>-0.02</td>
<td><strong>.036</strong></td>
<td>0.11</td>
<td>0</td>
<td>0.920</td>
</tr>
<tr>
<td>Inhibit delay (ms)</td>
<td>3</td>
<td>76</td>
<td>63</td>
<td>-0.53</td>
<td>-0.87</td>
<td>-0.19</td>
<td><strong>.002</strong></td>
<td>0.25</td>
<td>0</td>
<td>0.334</td>
</tr>
<tr>
<td>No-go reaction time (ms)</td>
<td>3</td>
<td>88</td>
<td>64</td>
<td>0.20</td>
<td>-0.12</td>
<td>0.53</td>
<td>.216</td>
<td>1.81</td>
<td>0</td>
<td>0.621</td>
</tr>
<tr>
<td><strong>Reasoning and problem solving</strong></td>
<td>12</td>
<td>224</td>
<td>301</td>
<td>0.43</td>
<td>-0.12</td>
<td>0.97</td>
<td><strong>.125</strong></td>
<td><strong>91.14</strong></td>
<td>88</td>
<td>0.889</td>
</tr>
<tr>
<td>Wisconsin Card Sorting Test (WCST)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conceptual level responses</td>
<td>4</td>
<td>81</td>
<td>155</td>
<td>-0.55</td>
<td>-0.83</td>
<td>-0.27</td>
<td>&lt;<strong>.001</strong></td>
<td>1.99</td>
<td>0</td>
<td>0.554</td>
</tr>
<tr>
<td>Failure to maintain set</td>
<td>3</td>
<td>57</td>
<td>52</td>
<td>0.63</td>
<td>-0.18</td>
<td>1.44</td>
<td>.127</td>
<td><strong>8.47</strong></td>
<td>76</td>
<td>0.391</td>
</tr>
<tr>
<td>N of categories</td>
<td>6</td>
<td>101</td>
<td>126</td>
<td>-0.42</td>
<td>-0.69</td>
<td>-0.15</td>
<td><strong>.002</strong></td>
<td>3.71</td>
<td>0</td>
<td>0.538</td>
</tr>
<tr>
<td>Non-perseverative errors</td>
<td>6</td>
<td>100</td>
<td>103</td>
<td>-0.20</td>
<td>-0.83</td>
<td>0.44</td>
<td>.543</td>
<td><strong>25.00</strong></td>
<td>80</td>
<td>0.222</td>
</tr>
<tr>
<td>Perseverative errors</td>
<td>11</td>
<td>203</td>
<td>280</td>
<td>0.30</td>
<td>-0.34</td>
<td>0.95</td>
<td>.360</td>
<td><strong>105.53</strong></td>
<td>91</td>
<td>0.751</td>
</tr>
<tr>
<td>Total errors</td>
<td>3</td>
<td>54</td>
<td>52</td>
<td>0.32</td>
<td>-0.60</td>
<td>1.25</td>
<td>.494</td>
<td><strong>11.15</strong></td>
<td>82</td>
<td>0.071</td>
</tr>
<tr>
<td><strong>Verbal learning and memory</strong></td>
<td>7</td>
<td>150</td>
<td>251</td>
<td>0.76</td>
<td>0.29</td>
<td>1.22</td>
<td><strong>.001</strong></td>
<td><strong>24.93</strong></td>
<td>76</td>
<td>0.355</td>
</tr>
<tr>
<td>California Verbal Learning Test for Children (CVLT-C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>t</td>
<td>Effect Size</td>
<td>p</td>
<td>d</td>
<td>Mean</td>
<td>SD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>------</td>
<td>-----</td>
<td>------</td>
<td>-------------</td>
<td>-------</td>
<td>------</td>
<td>------</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long delay free recall</td>
<td>3</td>
<td>85</td>
<td>105</td>
<td>-0.90</td>
<td>-1.65</td>
<td>-0.15</td>
<td>0.019</td>
<td>10.53d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recognition</td>
<td>3</td>
<td>89</td>
<td>154</td>
<td>-1.13</td>
<td>-2.40</td>
<td>0.15</td>
<td>0.084</td>
<td>33.60d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short delay free recall</td>
<td>3</td>
<td>85</td>
<td>105</td>
<td>-0.89</td>
<td>-1.37</td>
<td>-0.42</td>
<td>&lt; .001</td>
<td>4.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOMAL Facial Memory</td>
<td>3</td>
<td>49</td>
<td>56</td>
<td>-0.68</td>
<td>-1.39</td>
<td>0.03</td>
<td>0.062</td>
<td>6.15d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working memory</td>
<td>8</td>
<td>159</td>
<td>274</td>
<td>0.99</td>
<td>0.64</td>
<td>1.35</td>
<td>&lt; .001</td>
<td>18.29d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WISC Digit-span Forward</td>
<td>4</td>
<td>77</td>
<td>106</td>
<td>-0.79</td>
<td>-1.10</td>
<td>-0.48</td>
<td>&lt; .001</td>
<td>1.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed of processing</td>
<td>3</td>
<td>65</td>
<td>91</td>
<td>1.27</td>
<td>-0.12</td>
<td>2.65</td>
<td>0.074</td>
<td>27.51d</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual learning and memory</td>
<td>4</td>
<td>64</td>
<td>71</td>
<td>0.78</td>
<td>0.43</td>
<td>1.13</td>
<td>&lt; .001</td>
<td>2.79</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Statistically significant results are in bold. TOMAL = Test of Memory and Learning; WAIS = Wechsler Adult Intelligence Scale; WISC = Wechsler Intelligence Scale for Children.

- Effect sizes for the individual test measures are in the original test direction (i.e., an effect size larger than zero indicates that the measure is increased in cases, while an effect size smaller than zero indicates that the measure is decreased). Domains and global cognition positive effect size measures indicate cognitive impairment in cases, and are computed by averaging individual test measures.
- Global cognition score was calculated only in studies that evaluated 3 or more cognitive domains.
- P-value in Egger’s test for publication bias.

\[ P < .05 \]
References


<table>
<thead>
<tr>
<th></th>
<th>Reference</th>
</tr>
</thead>
</table>


Screening

Eligibility

Included

Identification

Records identified through database searching (n = 2,045)

Additional records identified through other sources (n = 780)

Records after duplicates removed (n = 1,909)

Records screened (n = 1,909)

Records excluded (n = 1,675)

Full-text articles excluded, with reasons (n = 210)

125 = Not children or not adolescent
4 = Did not include a control group
29 = Did not measure a neurocognitive function
23 = Not euthymic
7 = Insufficient data
13 = Authors contacted, no response
6 = Did not use a MATRICS test
2 = Other language
1 = Overlapping sample with included study

Full-text articles assessed for eligibility (n = 234)

Studies included in qualitative synthesis (n = 24)

Studies included in meta-analysis (N = 24)
Study ID Hedges's g (95% CI)
Voelbel, 2006 -0.66 (-1.44, 0.12)
Rucklidge, 2006 0.27 (-0.37, 0.92)
McClure, 2005 0.48 (-0.04, 1.00)
Kleinman, 2015 0.36 (-0.49, 1.22)
Udal, 2014 0.13 (-0.32, 0.59)
Udal, 2012a -0.49 (-0.96, -0.03)
Deveney, 2012 0.12 (-0.43, 0.67)
Kim, 2012 0.17 (-0.40, 0.73)
DelBello, 2004 0.36 (-0.50, 1.21)
Weathers, 2012 -0.02 (-0.66, 0.63)
Coelho, 2013 0.64 (0.03, 1.25)
Karakurt, 2013 0.85 (0.28, 1.42)
Overall 0.18 (-0.06, 0.43)
<table>
<thead>
<tr>
<th>Study</th>
<th>Hedges's g (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lera-Miguel, 2011</td>
<td>-0.87 (-1.60, -0.13)</td>
</tr>
<tr>
<td>Voelbel, 2006</td>
<td>-0.31 (-1.08, 0.45)</td>
</tr>
<tr>
<td>Rucklidge, 2006</td>
<td>0.14 (-0.49, 0.78)</td>
</tr>
<tr>
<td>Olvera, 2005</td>
<td>-0.77 (-1.62, 0.08)</td>
</tr>
<tr>
<td>Dickstein, 2004</td>
<td>0.40 (-0.21, 1.00)</td>
</tr>
<tr>
<td>Lera-Miguel, 2015</td>
<td>0.71 (0.08, 1.34)</td>
</tr>
<tr>
<td>Udal, 2012</td>
<td>0.57 (0.10, 1.03)</td>
</tr>
<tr>
<td>Franco, 2009</td>
<td>-0.19 (-0.99, 0.62)</td>
</tr>
<tr>
<td>Bedoya-Tovar, 2011</td>
<td>0.45 (-0.16, 1.07)</td>
</tr>
<tr>
<td>Coelho, 2013</td>
<td>0.22 (-0.37, 0.82)</td>
</tr>
<tr>
<td>Ray, 2012</td>
<td>3.73 (2.89, 4.57)</td>
</tr>
<tr>
<td>Karakurt, 2013</td>
<td>1.05 (0.46, 1.63)</td>
</tr>
<tr>
<td>Overall</td>
<td>0.43 (-0.12, 0.97)</td>
</tr>
</tbody>
</table>