MR. DIMOSTHENIS TSAPEKOS (Orcid ID : 0000-0002-1972-4813)

Article type : Review

Corresponding author mail id: dimosthenis.tsapekos@kcl.ac.uk

Title:

Cognitive enhancement interventions for people with bipolar disorder: a systematic review of methodological quality, treatment approaches, and outcomes.

Running title: Cognitive enhancement interventions for bipolar

Authors:

Dimosthenis Tsapekos[1], Benedetta Seccomandi[2], Tim Mantingh[1], Matteo Cella[2,3], Til Wykes[2,3] & Allan H. Young[1,3]

[1]Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK.

[2]Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK.

[3]South London & Maudsley NHS Foundation Trust, Maudsley Hospital, London, UK.

Abstract word count: 243

Text word count: 5,284

Acknowledgements

This paper represents independent research part-funded by the National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. DT would like to acknowledge financial support from the Onassis Foundation. TW would like to acknowledge her NIHR Senior Investigator award.

Disclosure

AHY reports paid lectures and advisory boards for the following companies with drugs used in affective and related disorders: AstraZeneca (AZ), Eli Lilly, Lundbeck, Sunovion, Servier, LivaNova, and Janssen. No shareholdings in pharmaceutical companies. Lead Investigator for Embolden Study (AZ), BCI Neuroplasticity study and Aripiprazole Mania Study. Investigator initiated studies from AZ, Eli Lilly, Lundbeck, Wyeth, Janssen. DT, BS, TM, MC, and TW report no conflict of interest for this work.

Abstract

Background: Patients with bipolar disorder (BD) suffer from cognitive deficits across several domains. The association between cognitive performance and psychosocial functioning has led to the emergence of cognition as a treatment target.

Objective: This study reviews the existing literature on cognitive enhancement interventions for people with BD, focusing on different treatment approaches and methodological quality.

Methods: We conducted a systematic search following the PRISMA guidelines. Sample characteristics and main outcomes for each study and treatment characteristics for each approach were extracted. Study quality was assessed using the Clinical Trials Assessment Measure (CTAM) and Cochrane Collaboration's Risk of Bias tool by independent raters.

Results: Eleven articles reporting data from seven original studies were identified encompassing 471 participants. Two treatment approaches were identified, cognitive and functional remediation. For controlled studies, methodological quality was modest (average CTAM score = 60.3), while the overall risk of bias was considered moderate. Beneficial effects on cognitive or functional outcomes were reported in the majority of studies (91%), but these findings were isolated and not replicated across studies. Key methodological limitations included small sample sizes, poor description of randomization process, high attrition rates, and participant exclusion from the analysis.

Conclusions: Findings are promising but preliminary. Quality studies were few and mostly underpowered. Heterogeneity in sample characteristics, outcome measures, and treatment approaches further limit the ability to generalize findings. Adequately powered trials are required to replicate initial findings, while moderators of treatment response and mechanisms of transfer need to be explored.

Keywords: bipolar disorder, cognition, functioning, cognitive remediation, functional remediation, methodological quality.

1. Introduction

Cognitive dysfunction is common in people with bipolar disorder (BD) and presents a profile qualitatively similar to that of people with schizophrenia, but less severe overall (1, 2). Impairment occurs in several cognitive domains with relevant meta-analyses reporting moderate to large deficits in attention, verbal learning and memory, and executive function (3, 4). These deficits are more pronounced during mood episodes but do persist and remain clinically significant after symptom remission and during euthymia (3, 5). What remains unclear is the relationship between illness progression and cognitive dysfunction for people with BD. Impairments may reflect neurodevelopmental factors, such as genetic susceptibility, and illness progression, such as the neurotoxic effects of recurrent mood episodes as well as chronic medication use (6). However, according to recent evidence only 40% of people with BD present significant deficits across multiple cognitive domains while another subgroup (30%) shows selective moderate impairments and the rest remain cognitively intact (7, 8). This heterogeneity poses serious methodological and clinical challenges when assessing the efficacy of interventions targeting cognitive impairment in this group of patients.

Cognitive interventions have emerged as a new treatment option to promote functional recovery of patients with BD. The association between cognition and function has been extensively explored, with most evidence suggesting significant contributions of cognitive impairment to reduced functional capacities (9, 10). Cognitive measures have been found to predict psychosocial functioning independently to residual depressive symptoms (11). Importantly, recent evidence from a large meta-analysis showed that cognitive performance, particularly verbal memory and executive function, is one of the strongest predictors of favourable employment outcomes (12). Despite this evidence, there is limited research on cognitive enhancement interventions for people with BD. For pro-cognitive pharmacological treatments, only preliminary evidence is reported from trials with severe methodological limitations (13). Findings on psychological treatments targeting cognitive impairment, such as cognitive remediation (CR), were also limited until recently. CR is a well-established intervention tackling cognitive and functional difficulties in people with schizophrenia, but CR approaches for people with BD have only emerged over the past few years (14). A meta-analysis using only the subgroups of participants with affective diagnoses from mixed sample studies shows that CR may have comparable benefits on cognition in affective and schizoaffective disorders as in schizophrenia (15). Subsequent studies have further explored the effects of CR and other interventions on cognitive and functional outcomes in people with BD.

The current report aims to review the evidence from studies examining cognitive enhancement interventions for people with BD in a systematic way. Tackling cognitive and functional difficulties is a new research field in BD. To consolidate a wider evaluation of its current status this review will emphasize two further aspects: a comparison of the intervention approaches assessed in people with BD and the methodological quality of the available studies. We believe that a critical reflection on the existing evidence quality will further appraise the reported efficacy of these interventions. Similarly, a comprehensive comparison of different treatment approaches targeting cognitive dysfunction may elucidate the clinical implications of previous findings and provide useful directions for improving future research in the field.

2. Methods

This systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses guidelines (PRISMA) (16). A protocol of the review was registered with the online database PROSPERO (registration number: CRD42018092130).

2.1 Data sources and search strategy

Systematic searches from the first available date to March 2019 were conducted in the following electronic databases: MEDLINE (PubMed), Embase, PsycINFO, Web of Science, and the Cochrane Collaboration Controlled Trials Register. The following search terms were used as keywords for the searches: (bipolar OR bd OR bpad OR manic-depress* OR manic depress*) AND (cognit* OR neurocognit* OR neuropsycholog*) AND (remediation OR rehabilitation OR training OR enhancement OR therapy). To identify any additional reports potentially missed, the reference lists of included studies, previous relevant reviews, and notable articles in the area were inspected.

2.2 Eligibility criteria

The following inclusion and exclusion criteria were used:

Participants: We considered studies involving adult participants (aged between 18-65) with a BD diagnosis according to the Diagnostic and Statistical Manual of Mental Disorders (17), the International Classification of Diseases (18), or the Research Diagnostic Criteria (19). Studies with mixed samples were only considered if participants with BD represented 40% or more of the sample.

Interventions: All interventions targeting any cognitive and/or any functional outcomes were considered. We also considered the combination of cognitive enhancement with other interventions (e.g., social cognitive training) if the cognitive training was a significant therapy component and accounted for a significant proportion of the therapy time (i.e., ≥ 40%).

Comparators: Eligible comparators included other active treatments not primarily targeting cognitive difficulties (e.g., psychoeducation, group support etc.), treatment-as-usual (TAU), or waiting list.

Outcomes: Only studies employing validated measures were considered.

Study design: All relevant studies were considered including randomized and quasi-randomized controlled trials, feasibility or pilot studies, and uncontrolled designs. We also considered sub-analyses and follow-up analyses reported in different manuscripts than the original study. Randomized controlled trials (RCTs) were prioritized in data synthesis and interpretation of the evidence.

2.3 Study selection

First the titles and then the abstracts of identified studies were screened by two authors independently (DT, BS). Full text reports were retrieved for all the articles considered eligible based on the first screening and were further reviewed for inclusion (DT, BS). Disagreements on the included studies were resolved based on consensus between the reviewers and, if required, after discussion with a third senior author (MC or TW).

2.4 Data extraction and synthesis

Two authors (DT, TM) independently extracted the following data from each article to collate relevant and comparable information from different reports: first author and year of publication, sample size and number of participants with BD, study design, demographic (age, gender, education) and clinical characteristics (premorbid IQ, BD type, illness duration, number of episodes, hospitalizations, current mood), type of intervention and control, assessment time-points and outcome measures, main findings, discontinuation rates, adverse events, and limitations. The following characteristics were extracted for each treatment approach: type and delivery method, setting, duration, main targets and core therapy components. Treatment outcomes are presented in a narrative synthesis, grouped by intervention type to enable comparisons and facilitate evidence interpretation.

2.5 Quality assessment and risk of bias

Methodological quality of controlled trials was assessed using the Clinical Trials Assessment Measure (CTAM) (20, 21). This is a 15-item scale assessing the methodology of studies examining psychological treatments. The features assessed by the CTAM include: sample characteristics, treatment allocation, comparators, assessment of outcome measures, active treatment description, and analysis, with quality ratings ranging between 0 and 100. Two reviewers (DT, TM) independently evaluated each report and any rating discrepancies were resolved by a senior author (MC). To ensure the accuracy of the ratings, and to avoid confusion between conduct quality and reporting quality, all corresponding authors were contacted and asked to examine the accuracy of our ratings. If necessary, amendments were made according to their feedback. The risk of bias for RCTs was also evaluated by two reviewers (DT, TM) using the Cochrane Collaboration's Risk of Bias tool (22).

3. Results

3.1 Study characteristics

The search and selection process are described in the PRISMA flowchart (Figure 1). Eleven articles reporting data from seven original studies were included in the review, assessing a total number of 471 participants: three were RCTs, two were non-

randomized controlled trials, and two were uncontrolled studies. All seven original studies were conducted by independent research groups and included unique sample sizes. The remaining four reports were follow-up or secondary analyses of the same RCT.

--- Figure 1 here ---

Three studies were conducted in the USA, and the remaining four in Europe. Sample sizes varied greatly, ranging from 12 to 239 participants and with a mean of 67. Participants across studies were primarily women (57%) and had an average age of 38.2 years. Five studies reported education attendance (13.4 years) and another five estimates of premorbid IQ (108.7). Most participants had a diagnosis of BD type I (64.7%) and an average illness duration of 14.7 years (range: 7-28). Participants had experienced 8.9 mood episodes and had been admitted to the hospital 2.9 times on average. All studies recruited outpatients. Table 1 displays the characteristics of all included reports.

--- Table 1 here ---

3.2 Intervention characteristics

The main characteristics and aims for each intervention are described in Table 2. Other than traditional CR approaches, a novel remediation approach for people with BD was identified, functional remediation (FR). Out of the seven original studies, five investigated CR and six FR approaches. None were combined with another active treatment. Of the controlled studies, three used treatment-as-usual (TAU) as a comparator, one CR trial used computer games as an active control condition, and five FR reports had both TAU and an active intervention (i.e., psychoeducation).

Most studies (63%) delivered therapy through pen-and-paper activities; the remaining used various types of computer software. In terms of setting, three interventions were delivered individually, while the rest employed a group or a combination format with small group sizes (4-8 participants) and one or more group facilitators. One study also invited caregivers to attend psychoeducational group sessions. Therapy duration varied from 12 to 24 weeks with an average of 16.2 weeks. The target for treatment sessions ranged from 12 to 70 and had an average of 23.5, but most interventions supplemented group or individual sessions with learning materials about BD and homework exercises or tasks. However, only two CR programmes delivered highly intensive training with multiple sessions weekly. These were the only ones not using a therapist to support or guide participants during their cognitive training (23, 24).

CR and FR approaches presented a similar adherence profile. Mean treatment discontinuation rates were 28.7% for CR and 28.6% for FR. The main reasons for discontinuing therapy were workload and time requirements, lack of motivation to engage with the therapy, and changes in mental health status. CR satisfaction was assessed only in one trial where participants in the active therapy group reported moderate-to-high ratings using a feedback survey (23). Satisfaction for FR was examined only by an uncontrolled study with participants highly rating group and individual sessions (25). CR and FR were tolerated well: adverse events were scarce and equally balanced between active and control groups. Exacerbation of affective symptoms was reported from 19 participants across five studies, with 10 receiving an active therapy and nine the control condition.

--- Table 2 here ---

3.3 Methodological quality

Table 3a displays the CTAM ratings for the included controlled studies. The mean CTAM score was 60.3 (SD 15.1) with a range of 29 to 74. All corresponding authors were contacted, and three replied with their feedback: two agreed with our initial rating and one suggested an amendment. The total CTAM increase was only three points. Only four studies (44%) had a score above the cut-off of 65 (21), indicating low susceptibility to bias. The primary problem was low sample sizes (44%), unblinded assessments (77%), not applying intention-to-treat (ITT) analysis (66%), lack of treatment protocols (33%) and fidelity assessments (100%). Most studies did allocate participants randomly (78%), but this randomization process was commonly not independent to the study team (88%). Another problem was high attrition rates which ranged from 12.8% to 38.4% (mean: 26.6%) for active treatment groups and from 4.3% to 30.5% (mean: 17.1%) for control groups. Only one study reported an overall drop-out rate lower than 15%.

Table 3b shows the risk of bias ratings for included RCTs and RCT sub-analyses. The overall risk of bias was considered moderate according to the Cochrane tool which corresponds with the mean CTAM rating for the included reports.

--- Tables 3a and 3b here ---

3.4 Intervention outcomes

The largest CR trial to date (N=75) reported significant improvement for the treatment group in a global cognition score (Cohen's d=0.8), as well as moderate-to-large effects in visual-spatial learning (Cohen's d=0.92) and processing speed (Cohen's d=0.65) (23).

Though cognitive improvement was correlated with changes in the primary functional outcome across the sample, a significant CR effect over computer control was not detected for community functioning. An earlier smaller trial (N=46) showed no beneficial effects of CR over standard treatment on verbal memory or any other secondary outcomes (26). CR was only associated with improvements in verbal fluency (Cohen's d=0.25) and quality of life (Cohen's d=0.75) at the 6-month follow-up. Two non-randomized studies have presented moderate effects of CR on cognitive outcomes over control conditions. Preiss et al. (24) delivered computerized CR to 15 depressed outpatients with unipolar or bipolar disorder and reported post-treatment reduction in attentional and executive difficulties compared to 16 age- and diagnosis-matched controls. In a study where participants were self-selected for CR or TAU, authors reported improvements in problem solving and working memory for the treatment group (27). The earliest study to examine CR for people with BD (N=18) reported improvement in an observer-rated measure of executive dysfunction, and measures of occupational and psychosocial functioning after 14 CR sessions (28). In total, four out of five CR studies (80%) showed significant improvement in one or more cognitive outcomes with effect sizes ranging from small to large (0.35-0.95). For functional outcomes, only two studies (40%) reported significant changes at post-treatment or follow-up. Although all studies explored the association between cognitive and functional changes, a significant correlation was not detected specifically for the CR group in any of those.

FR was primarily investigated in a large randomized trial (N=239) which allocated participants to FR (N=77), psychoeducation (PE; N=82) or TAU (N=80) (29). The post-treatment effect of FR on psychosocial functioning was small-to-moderate versus TAU (Cohen's d=0.3), but a significant difference was not found over the PE group. For cognitive outcomes, no significant benefits were reported for FR over PE or TAU. A follow-up analysis one year after baseline (N=172) showed that FR maintained a small effect on functioning over TAU (Cohen's d=0.18), and additionally demonstrated an improvement in verbal memory for the FR group only (30). Several post-hoc analyses used data from this original trial to further explore the effect of FR on cognition and functioning. Bonnin et al (31) analysed a subgroup of participants presenting objective cognitive impairment (i.e., score of 2 SDs or more below the normative mean in one or more measures; N=188) and reported a small effect of FR over TAU in verbal memory (Cohen's d=0.3), while a small effect over control was found for psychosocial functioning (Cohen's d=0.2). FR was not superior to PE in any of these measures. A small sub-analysis of the 53 participants with a bipolar type-II diagnosis did not replicate any of the previous effects of FR on cognition or functioning over TAU, showing only a reduction in depressive symptoms (32). However, a larger sub-analysis considering participants with subthreshold depressive symptoms (N=99) did report a significant functional improvement over TAU (Cohen's d=1.12) and PE (Cohen's d=0.37) which were maintained at 1-year follow-up (33). No differences were found for any of the cognitive measures. Recently, a small uncontrolled study (N=12) described improved psychosocial functioning following a brief group FR focused on personalized goals (25). Overall, five out of six FR studies (83.4%) reported significant benefits for psychosocial functioning with a moderate mean effect (Cohen's d=0.45) using the same clinician-rated measure (Functional Assessment Short Test [FAST]) (34). Cognitive benefits were only demonstrated for verbal memory in two reports

(33.3%) with a small mean effect (Cohen's d=0.25). A significant association between cognitive and functional changes specifically for the FR group was only detected in one study (30).

The association of methodological quality with treatment outcomes was unclear for the included studies. CTAM scores ranged from 29 to 74 for studies reporting significant changes in one or more outcomes. The same range for studies with no significant effects was 62 to 70. There was a negative but not significant correlation (Spearman's r=-0.18) between CTAM scores and effect sizes reported for cognitive and functional outcomes. Sample size possibly was a key variable for this relationship: larger studies were more rigorous (Spearman's r=0.74, p=0.24) and associated with reporting smaller effect sizes (Spearman's r=-0.71, p=0.31).

4. Discussion

Treating cognitive impairment in people with BD may benefit psychosocial difficulties and improve occupational outcomes (11, 35). This review evaluated the evidence on cognitive enhancement interventions for people with BD focusing on the different treatment paradigms available and the methodological quality of the included studies. Two interventions have been considered for people with BD, cognitive and functional remediation. CR approaches were heterogeneous in terms of delivery method, therapy components and goals. Methodology varied significantly among the included reports and quality ratings were moderate on average. Cognitive enhancement interventions showed positive findings for cognitive and functional outcomes, but these were not robust and can only be considered preliminary, which is in line with recent relevant reviews(13, 36, 37).

4.1 Cognitive remediation

CR comprised various treatment approaches with common characteristics but also significant differences. Literature on people with schizophrenia has divided CR into two paradigms: restorative and compensatory (38). Restorative approaches have emphasized intensive and repetitive training with increasing levels of difficulty (39), while compensatory ones combine cognitive training with strategy learning which is considered the key to transfer any cognitive improvement into functional outcomes (40). For studies considered in this review, the main differences between restorative and compensatory approaches were the role of the therapist and the number of sessions. Therapists adopted more active roles in compensatory CR approaches aiming to emphasize strategy use and to bridge individual or group training with daily-life activities (26). For restorative programmes, the primary focus was massed computer training with progressively more difficult tasks and multiple sessions per week, while it was not clearly reported if any guidance or support was provided by a therapist (23).

CR was accepted well and was associated with some cognitive improvements in the studies considered. However, evidence was inconsistent, with some studies reporting beneficial treatment effects and others no significant changes. This may reflect the methodological limitations of most studies, particularly small sample sizes with inadequate power to detect treatment effects. Significant variability in the CR approaches (e.g., every study used a different computer software) might have contributed to that. Studies recruited samples with different characteristics at baseline which might have affected treatment outcomes. For example,

Lewandowski et al. (23) recruited patients with history of psychosis which has been associated with greater cognitive impairments and larger potential for improvement (41).

Overall, results are in line with previous reviews in this population. Anaya et al. (15) conducted a meta-analysis of 18 CR studies having a significant proportion of participants with schizo-affective or affective disorders in their sample (N=300; 35%) and reported a pooled mean effect size of 0.36 for cognitive outcomes. As in the current review, authors described significant methodological limitations such as the uncontrolled design of several studies and the heterogeneity of CR approaches. A more recent CR review for patients with affective psychosis reported moderate-to-large effect sizes (Cohen's d=0.36-0.94), but findings were not replicated for most domains across studies (42). Examining the effect of CR on different cognitive domains in the studies considered, we found minimal consistency in post-treatment benefits since most significant findings did not replicate. The only domain showing improvement in several studies was executive functioning, such as problem solving and set shifting. This is an encouraging finding given that impairments in executive functions have been proposed as a cognitive endophenotype of BD (43).

The scarcity of high quality and adequately powered trials in people with BD limits the comparison to people with schizophrenia where a robust body of evidence suggests at least moderate effect sizes (44). Patients with schizophrenia show more pronounced cognitive impairment and are more likely to benefit following CR than patients with affective diagnoses since they are less susceptible to ceiling effects in cognitive measures (45). This baseline discrepancy might disappear though, if only BD patients with a history of psychosis or with objective cognitive impairments are considered. This was supported by a recent meta-regression of five RCTs (N=300) examining CR in mixed samples with severe mental illness (SMI) including schizophrenia, schizo-affective and affective disorders. Authors confirmed lower baseline cognitive performance is associated with greater response to CR, but this was independent of diagnosis type (46).

To date, there have not been any trials comparing the effects of CR among different diagnostic groups. However, it is worth commenting on a significant amount of CR studies identified from our search (n=12) with SMI samples including a percentage of patients with BD (N with BD=162; mean proportion: 26%; range: 6-34%). A randomized trial of individual CR combined with group discussions described significant improvement in processing speed and verbal memory for the treatment group over computerized control (47). Participants in this group were employed at a greater rate compared to control at the 12-month follow-up. In a recent RCT with 24% of the sample diagnosed with BD, CR was associated with significant improvement in executive functioning, while participants in the treatment group reported better work outcomes over the next two years (48). Although none of these studies presented results separately for different diagnostic groups, cognitive and functional enhancement were reported for participants across diagnoses. Thus, diagnosis may be less important to treatment response than other individual factors such as objective cognitive difficulties. Finally, mixed sample studies had several methodological limitations and examined variable CR approaches, similar to the studies considered in this review.

4.2 Functional remediation

FR was developed specifically to address difficulties for people with BD and overlaps greatly with compensatory CR in terms of treatment mechanisms. Both approaches attempt to teach adaptive strategies and techniques to overcome cognitive and functional difficulties (49). Unlike CR studies, where multiple approaches were examined, all FR studies involved the same therapy protocol and all, but one uncontrolled study, were conducted by the same research group. This protocol emphasized education about cognitive deficits and their impact on everyday life, as well as training of functional skills using role-playing and real-life exercises. Results from these studies suggest greater improvements for FR over TAU in overall psychosocial functioning, which was assessed consistently across studies using a recommended measure (FAST). Benefits were particularly demonstrated for the domains of autonomy, interpersonal relationships, and occupational functioning.

Although still not replicated by other studies, results from this group are in line with findings from schizophrenia research. A previous meta-analysis considering the effects of CR on functional outcomes reported a moderate mean effect size (Cohen's d=0.42) which was durable over time (44). Benefits for functional outcomes were linked to compensatory remediation approaches, like FR, over restorative ones. However, what FR studies have yet to examine is the efficacy on 'hard' occupational outcomes, such as employment rates and wages over long follow-up periods, which have been reported from similar therapeutic approaches in SMI studies (48, 50). For cognitive outcomes, verbal memory was the only domain found to improve following FR, and importantly the only domain associated with changes in psychosocial functioning over time (30). Impairment in this domain is a strong predictor of psychosocial functioning in BD and any changes would normally affect functional performance (51).

4.3 Limitations

There were several factors limiting the studies considered in this review.

Methodological quality: The mean CTAM score was 60.3 which is below the threshold of high methodological quality according to Wykes et al. (21). This might indicate that treatment effects have been overestimated due to poor methodological quality. Previous reviews using this measure have reported comparable mean CTAM scores, 61.2 for trials examining cognitive behavioral therapy (CBT) (21) and 57.4 for trials assessing CR interventions (44). The review of CBT detected an association between methodological rigor and reported effect sizes, while methodological ratings did not correlate with therapy effects for CR trials. Unlike these reviews where a large number of studies were considered, the small number of studies included in our review potentially limits the capacity to further explore the relationship between quality ratings and changes in outcome measures. CR for people with BD is an emerging research field and most of the studies to date were feasibility or pilot trials which commonly involve methodological shortcomings and cannot reliably estimate effect sizes due to small sample sizes. Regarding measure selection, most studies did not have a clear a priori defined hierarchy of primary and secondary outcomes. Statistical analysis was limited due to the violation of the ITT principle, the lack of imputation methods to appropriately handle missing data (e.g., mixed-models approach), and the absence of control for multiple testing (e.g., Bonferroni correction). Another limitation was the way treatment effect was estimated in some studies which chose to report within-group effect sizes despite including a control

condition. Within-group effects can overestimate the impact of an intervention since changes in the control group are not considered.

Heterogeneity: Great variability in sample characteristics, cognitive measures, and treatment approaches was present across studies. Only two reports included objective cognitive screening. This led to cognitively heterogeneous study samples since not every patient with BD had cognitive deficits (52). Such samples involve the risk of underestimating the efficacy of an intervention due to lack of impairments in a significant proportion of the sample and, thus, allowing little room for improvement following therapy. Regarding clinical characteristics, not every study recruited participants who were free of acute affective symptoms which may have hindered therapy delivery and reliability of cognitive assessments. Mood symptoms can limit the impact of an intervention on cognition, while cognitive improvement might be the product of changes in mood outcomes (53). Although most studies used objective and validated measures, cognitive performance was examined differently across studies and this limits the comparability between findings. Studies adopted a variety of measures to assess individual cognitive domains and did not report composite cognitive scores to assess global cognition, with the exception of one trial using the MATRICS Consensus Cognitive Battery (MCCB) (54). None of the studies considered in this review used the cognitive battery recommended by the International Society for Bipolar Disorders (55). Heterogeneity was also evident in CR approaches, especially in regard to duration and intensity, delivery method, and goal setting. These differences may limit the generalization of post-treatment improvements across different CR programmes. Variability in the included studies is described in Table 4.

Durability of outcomes: Most studies in this review included only short follow-up periods, with a mean of 25.5 weeks, while several did not conduct assessments after the end of therapy. Hence, it remains unclear if any post-therapy effects are durable over time. Another reason longer follow-up periods are required is to examine the association between cognitive and functional change. According to Miskowiak et al. (55), there is a time-lag between cognitive changes and improvement in functional competence. Trials need to allow follow-up periods of at least three months following therapy completion to explore this association. For example, studies with mixed SMI samples examining occupational outcomes after CR involved long follow-up periods, typically over one year (47, 48). In people with BD, the only study reporting a significant correlation between improvement in verbal memory and functional changes over time was the one with a 12-month follow-up period (30).

--- Table 4 here ---

4.4 Future research directions

An international task force has recently considered a series of methodological challenges for cognition trials in BD such as inclusion criteria, selection of outcomes, and statistical issues (55). Experts provided numerous consensus-based recommendations for future trials to improve their sensitivity in detecting treatment effects. Recommendations included enriching study samples for cognitive impairment by screening potential participants with objective cognitive measures, recruiting patients without acute mood symptoms and selecting a composite score as primary cognitive outcomes. Regarding psychological

interventions targeting cognition, appropriately designed and rigorous randomized trials are required to address current methodological limitations. Future trials should involve larger sample sizes to estimate efficacy of these interventions with adequate power, homogeneity in sample characteristics and outcome measures to enable generalisability and comparability of findings, and precise statistical analyses to avoid any inflation of treatment effects. Currently, there are some registered trials testing different CR approaches in patients with BD (56-58).

Larger trials may allow the exploration of mechanisms and patients' characteristics associated with treatment response. Factors such as age (59), pre-treatment cognitive performance (46), and cognitive reserve (23) are candidate moderators of CR response based on theory and some initial evidence from affective populations. The role of these factors in CR for people with BD needs to be further explored to improve treatment outcomes. Modern statistical techniques, such as machine learning, can refine the search for these underlying treatment mechanisms even in smaller samples (60). Additional moderating factors might include illness duration, number of mood episodes, and diagnostic subtype. The role of anxiety symptoms is also worth exploring as some evidence indicates an association with CR response for people with BD (28). Identifying response moderators may allow the development of personalized CR approaches for people with BD, tailored to the specific difficulties experienced from particular patient subgroups and adapted to different cognitive profiles. Recent findings point to distinct cognitive subgroups in terms of strengths and deficits in executive functions which might indicate a variability of cognitive phenotypes in BD and potentially different therapy requirements (61).

Another important area is the association between cognitive improvement and changes in psychosocial functioning. Only minimal findings have been reported about the translation of cognitive into functional gains in studies including people with BD (30, 62), potentially due to the limitations previously discussed. Future trials should examine the long-term benefits in daily life outcomes (e.g., work) and the transfer mechanisms connecting therapy with these functional improvements, as previously demonstrated for people with schizophrenia (63). In that context, future research should investigate the therapy components mostly associated with post-treatment cognitive and functional changes (64) since no indications currently exist of how therapy duration or delivery method may affect outcomes in people with BD.

Similarly, only few of the studies considered in this review explored the satisfaction rates among patients receiving therapy. Taking 15% as a minimum attrition threshold to ensure validity of the findings (44), both treatment discontinuation and trial drop-out were increased across the included studies suggesting low acceptability and showing that refinements are required in therapy protocols. In addition, none of the studies examined participant satisfaction in relation to treatment outcomes. Perceiving an intervention as beneficial can be critical for the overall treatment effect in the long-term (65). Future trials in BD need to address these issues by examining participant feedback for the therapy (66) and by considering attainment of personal goals as a secondary outcome to evaluate intervention utility from the patient's perspective (67).

Finally, a key target for future trials should be the identification of neurobiological markers associated with cognitive improvement. Neuroimaging evidence indicate aberrant frontal activation as an underlying factor for cognitive impairment in BD (68, 69) . Meta-analyses in people with schizophrenia have associated CR interventions with enhanced frontal activations suggesting this might be a reliable biomarker of cognitive improvement (70, 71). Only a few studies have examined potential changes in neural activation following CR in people with BD, with findings supporting the model of hypofrontality and that modulating activity in frontal areas, particularly in the dorsolateral prefrontal cortex, potentially underlies cognitive improvement (72, 73). Identifying neuroimaging markers of response to CR would be an important step to establish better understanding for the neurobiology of declined cognition in BD and to adapt the targets of cognitive training in future interventions.

5. Conclusions

Studies considered in this review suggest that cognitive enhancement interventions may benefit cognitive and functional outcomes, but the evidence is far from conclusive. The association between cognitive and functional improvement is unclear, while the mechanisms of this transfer have not been explored. Addressing the methodological limitations of previous studies is necessary to evaluate the efficacy of CR and FR in a reliable and accurate way. Potential moderators of response should be identified in order to develop more efficient and personalized treatment programmes.

References

1. Barch DM. Neuropsychological abnormalities in schizophrenia and major mood disorders: Similarities and differences. Current Psychiatry Reports. 2009;11(4):313-19.

2. Bora E, Yucel M, Pantelis C. Cognitive functioning in schizophrenia, schizoaffective disorder and affective psychoses: meta-analytic study. Br J Psychiatry. 2009;195(6):475-82.

3. Bourne C, Aydemir O, Balanza-Martinez V, Bora E, Brissos S, Cavanagh JT, et al. Neuropsychological testing of cognitive impairment in euthymic bipolar disorder: an individual patient data meta-analysis. Acta Psychiatr Scand. 2013;128(3):149-62.

4. Torres IJ, Boudreau VG, Yatham LN. Neuropsychological functioning in euthymic bipolar disorder: a meta-analysis. Acta Psychiatrica Scandinavica. 2007;116(s434):17-26.

5. Demmo C, Lagerberg TV, Kvitland LR, Aminoff SR, Hellvin T, Simonsen C, et al. Neurocognitive functioning, clinical course and functional outcome in first-treatment bipolar I disorder patients with and without clinical relapse: A 1-year follow-up study. Bipolar Disord. 2018;20(3):228-37.

6. Bora E, Ozerdem A. Meta-analysis of longitudinal studies of cognition in bipolar disorder: comparison with healthy controls and schizophrenia. Psychol Med. 2017;47(16):2753-66.

7. Burdick KE, Russo M, Frangou S, Mahon K, Braga RJ, Shanahan M, et al. Empirical evidence for discrete neurocognitive subgroups in bipolar disorder: clinical implications. Psychol Med. 2014;44(14):3083-96.

8. Bora E, Hidiroglu C, Ozerdem A, Kacar OF, Sarisoy G, Civil Arslan F, et al. Executive dysfunction and cognitive subgroups in a large sample of euthymic patients with bipolar disorder. Eur Neuropsychopharmacol. 2016;26(8):1338-47.
9. Depp CA, Mausbach BT, Harmell AL, Savla GN, Bowie CR, Harvey PD, et al. Meta-analysis of the association between cognitive abilities and everyday functioning in bipolar disorder. Bipolar Disord. 2012;14(3):217-26.
10. Martino DJ, Marengo E, Igoa A, Scapola M, Ais ED, Perinot L, et al. Neurocognitive and symptomatic predictors of functional outcome in bipolar disorders: a prospective 1 year follow-up study. J Affect Disord. 2009;116(1-2):37-42.
11. Roux P, Raust A, Cannavo A-S, Aubin V, Aouizerate B, Azorin J-M, et al. Associations between residual depressive symptoms, cognition, and functioning in patients with euthymic bipolar disorder: results from the FACE-BD cohort. The British Journal of Psychiatry. 2018;211(6):381-7.
12. Tse S, Chan S, Ng KL, Yatham LN. Meta-analysis of predictors of favorable employment outcomes among individuals with bipolar disorder. Bipolar Disord. 2014;16(3):217-29.
13. Miskowiak KW, Carvalho AF, Vieta E, Kessing LV. Cognitive enhancement treatments for bipolar disorder: A systematic review and methodological recommendations. Eur Neuropsychopharmacol. 2016;26(10):1541-61.
14. Bowie CR, Gupta M, Holshausen K. Cognitive Remediation Therapy for Mood Disorders: Rationale, Early Evidence, and Future Directions. The Canadian Journal of Psychiatry. 2013;58(6):319-25.
15. Anaya C, Aran AM, Ayuso-Mateos JL, Wykes T, Vieta E, Scott J. A systematic review of cognitive remediation for schizo-affective and affective disorders. Journal of Affective Disorders. 2012;142(1):13-21.
16. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS medicine. 2009;6(7):e1000097.
17. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5®): American Psychiatric Pub; 2013.
18. World Health Organization. International statistical classification of diseases and related health problems (11th Revision). Retrieved from https://icd.who.int/browse11/l-m/en2018.
19. Spitzer RL EJ, Robins E. Research diagnostic criteria: rationale and reliability. Archives of general psychiatry. 1978;35(6):773-82.
20. Tarrier N, Wykes T. Is there evidence that cognitive behaviour therapy is an effective treatment for schizophrenia? A cautious or cautionary tale? Behav Res Ther. 2004;42(12):1377-401.
21. Wykes T, Steel C, Everitt B, Tarrier N. Cognitive behavior therapy for schizophrenia: effect sizes, clinical models, and methodological rigor. Schizophrenia bulletin. 2008;34(3):523-37.
22. Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions: John Wiley & Sons; 2011.
23. Lewandowski KE, Sperry SH, Cohen BM, Norris LA, Fitzmaurice GM, Ongur D, et al. Treatment to enhance cognition in bipolar disorder (TREC-BD): Efficacy of a randomized controlled trial of cognitive remediation versus active control. Journal of Clinical Psychiatry. 2017;78(9):e1242-e9.

24. Preiss M, Shatil E, Cermakova R, Cimermanova D, Ram I. Personalized cognitive training in unipolar and bipolar disorder: A study of cognitive functioning. Frontiers in Human Neuroscience. 2013;14-23.
25. Zyto S, Jabben N, Schulte PFJ, Regeer BJ, Kupka RW. A pilot study of a combined group and individual functional remediation program for patients with bipolar I disorder. Journal of Affective Disorders. 2016;194:9-15.
26. Demant KM, Vinberg M, Kessing LV, Miskowiak KW. Effects of Short-Term Cognitive Remediation on Cognitive Dysfunction in Partially or Fully Remitted Individuals with Bipolar Disorder: Results of a Randomised Controlled Trial. Plos One. 2015;10(6):17-24.
27. Veeh J, Kopf J, Kittel-Schneider S, Deckert J, Reif A. Cognitive remediation for bipolar patients with objective cognitive impairment: a naturalistic study. International Journal of Bipolar Disorders. 2017;5(1):8-20.
28. Deckersbach T, Nierenberg AA, Kessler R, Lund HG, Ametrano RM, Sachs G, et al. Cognitive rehabilitation for bipolar disorder: An open trial for employed patients with residual depressive symptoms. CNS Neuroscience and Therapeutics. 2010;16(5):298-307.
29. Torrent C, Bonnin CD, Martínez-Arán A, Valle J, Amann BL, González-Pinto A, et al. Efficacy of Functional Remediation in Bipolar Disorder: A Multicenter Randomized Controlled Study. American Journal of Psychiatry. 2013;170(8):852-9.
30. Bonnin C, Torrent C, Arango C, Amann B, Sole B, Gonzalez-Pinto A, et al. Functional remediation in bipolar disorder: 1-year follow-up of neurocognitive and functional outcome. The British Journal of Psychiatry. 2016;208(1):87-93.
31. Bonnin C, Reinares M, Martinez-Aran A, Balanza-Martinez V, Sole B, Torrent C, et al. Effects of functional remediation on neurocognitively impaired bipolar patients: Enhancement of verbal memory. Psychological Medicine. 2016;46(2):291-301.
32. Solé B, Bonnin CM, Mayoral M, Amann BL, Torres I, González-Pinto A, et al. Functional remediation for patients with bipolar II disorder: Improvement of functioning and subsyndromal symptoms. European Neuropsychopharmacology. 2015;25(2):257-64.
33. Sanchez-Moreno J, Bonnín C, González-Pinto A, Amann BL, Solé B, Balanzá-Martínez V, et al. Do patients with bipolar disorder and subsyndromal symptoms benefit from functional remediation? A 12-month follow-up study. European Neuropsychopharmacology. 2017;27(4):350-9.
34. Rosa AR, Sánchez-Moreno J, Martínez-Aran A, Salamero M, Torrent C, Reinares M, et al. Validity and reliability of the Functioning Assessment Short Test (FAST) in bipolar disorder. Clinical Practice and Epidemiology in Mental Health. 2007;3(1):5.
35. O'Donnell LA, Deldin PJ, Grogan-Kaylor A, McInnis MG, Weintraub J, Ryan KA, et al. Depression and executive functioning deficits predict poor occupational functioning in a large longitudinal sample with bipolar disorder. J Affect Disord. 2017;215:135-42.
36. Sole B, Jimenez E, Torrent C, Reinares M, Bonnin CDM, Torres I, et al. Cognitive Impairment in Bipolar Disorder: Treatment and Prevention Strategies. Int J Neuropsychopharmacol. 2017;20(8):670-80.
37. Kim EJ, Bahk Y-C, Oh H, Lee W-H, Lee J-S, Choi K-H. Current Status of Cognitive Remediation for Psychiatric Disorders: A Review. Frontiers in psychiatry. 2018;9:461.
38. Medalia A, Choi J. Cognitive Remediation in Schizophrenia. Neuropsychology Review. 2009;19(3):353.

39. Fisher M, Holland C, Merzenich MM, Vinogradov S. Using neuroplasticity-based auditory training to improve verbal memory in schizophrenia. The American journal of psychiatry. 2009;166(7):805-11.
40. Wykes T, Spaulding WD. Thinking about the future cognitive remediation therapy--what works and could we do better? Schizophr Bull. 2011;37 Suppl 2:S80-90.
41. Bowie CR, Best MW, Depp C, Mausbach BT, Patterson TL, Pulver AE, et al. Cognitive and functional deficits in bipolar disorder and schizophrenia as a function of the presence and history of psychosis. Bipolar Disord. 2018;20(7):604-13.
42. Biagianti B, Merchant J, Brambilla P, Lewandowski KE. The effects of cognitive remediation in patients with affective psychosis: A systematic review: Special Section on "Translational and Neuroscience Studies in Affective Disorders". Section Editor, Maria Nobile MD, PhD. This Section of JAD focuses on the relevance of translational and neuroscience studies in providing a better understanding of the neural basis of affective disorders. The main aim is to briefly summaries relevant research findings in clinical neuroscience with particular regards to specific innovative topics in mood and anxiety disorders. Journal of Affective Disorders. 2019;255:188-94.
43. Bora E, Yucel M, Pantelis C. Cognitive endophenotypes of bipolar disorder: a meta-analysis of neuropsychological deficits in euthymic patients and their first-degree relatives. J Affect Disord. 2009;113(1-2):1-20.
44. Wykes T, Huddy V, Cellard C, McGurk SR, Czobor P. A meta-analysis of cognitive remediation for schizophrenia: methodology and effect sizes. American Journal of Psychiatry. 2011;168(5):472-85.
45. Twamley EW, Burton CZ, Vella L. Compensatory cognitive training for psychosis: who benefits? Who stays in treatment? Schizophrenia bulletin. 2011;37 Suppl 2:S55-62.
46. DeTore NR, Mueser KT, Byrd JA, McGurk SR. Cognitive functioning as a predictor of response to comprehensive cognitive remediation. Journal of Psychiatric Research. 2019;113:117-24.
47. Lindenmayer J-P, McGurk SR, Mueser KT, Khan A, Wance D, Hoffman L, et al. A Randomized Controlled Trial of Cognitive Remediation Among Inpatients With Persistent Mental Illness. Psychiatric Services. 2008;59(3):241-7.
48. McGurk SR, Mueser KT, Xie H, Welsh J, Kaiser S, Drake RE, et al. Cognitive Enhancement Treatment for People With Mental Illness Who Do Not Respond to Supported Employment: A Randomized Controlled Trial. American Journal of Psychiatry. 2015;172(9):852-61.
49. Martínez-Arán A, Torrent C, Solé B, Bonnín CM, Rosa AR, Sánchez-Moreno J, et al. Functional remediation for bipolar disorder. Clinical practice and epidemiology in mental health: CP & EMH. 2011;7:112-6.
50. Bowie CR, Grossman M, Gupta M, Holshausen K, Best MW. Action-based cognitive remediation for individuals with serious mental illnesses: Effects of real-world simulations and goal setting on functional and vocational outcomes. Psychiatr Rehabil J. 2017;40(1):53-60.
51. Bonnin CM, Martinez-Aran A, Torrent C, Pacchiarotti I, Rosa AR, Franco C, et al. Clinical and neurocognitive predictors of functional outcome in bipolar euthymic patients: a long-term, follow-up study. J Affect Disord. 2010;121(1-2):156-60.

52. Jensen JH, Knorr U, Vinberg M, Kessing LV, Miskowiak KW. Discrete neurocognitive subgroups in fully or partially remitted bipolar disorder: Associations with functional abilities. J Affect Disord. 2016;205:378-86.
53. Burdick KE, Braga RJ, Nnadi CU, Shaya Y, Stearns WH, Malhotra AK. Placebo-controlled adjunctive trial of pramipexole in patients with bipolar disorder: targeting cognitive dysfunction. The Journal of clinical psychiatry. 2012;73(1):103-12.
54. Burdick KE, Ketter TA, Goldberg JF, Calabrese JR. Assessing cognitive function in bipolar disorder: challenges and recommendations for clinical trial design. J Clin Psychiatry. 2015;76(3):e342-50.
55. Miskowiak KW, Burdick KE, Martinez-Aran A, Bonnin CM, Bowie CR, Carvalho AF, et al. Methodological recommendations for cognition trials in bipolar disorder by the International Society for Bipolar Disorders Targeting Cognition Task Force. Bipolar Disord. 2017;19(8):614-26.
56. Strawbridge R, Fish J, Halari R, Hodsoll J, Reeder C, Macritchie K, et al. The Cognitive Remediation in Bipolar (CRiB) pilot study: study protocol for a randomised controlled trial. Trials. 2016;17:9.
57. Gomes BC, Rocca CC, Belizario GO, Lafer B. Cognitive-behavioral rehabilitation vs. treatment as usual for bipolar patients: study protocol for a randomized controlled trial. Trials. 2017;18(1):142.
58. Ott CV, Vinberg M, Bowie CR, Christensen EM, Knudsen GM, Kessing LV, et al. Effect of action-based cognitive remediation on cognition and neural activity in bipolar disorder: study protocol for a randomized controlled trial. Trials. 2018;19(1):487.
59. Thomas KR, Puig O, Twamley EW. Age as a moderator of change following compensatory cognitive training in individuals with severe mental illnesses. Psychiatr Rehabil J. 2017;40(1):70-8.
60. Ramsay IS, Ma S, Fisher M, Loewy RL, Ragland JD, Niendam T, et al. Model selection and prediction of outcomes in recent onset schizophrenia patients who undergo cognitive training. Schizophr Res Cogn. 2018;11:1-5.
61. Kollmann B, Yuen K, Scholz V, Wessa M. Cognitive variability in bipolar I disorder: A cluster-analytic approach informed by resting-state data. Neuropharmacology. 2019.
62. Medalia A, Dorn H, Watras-Gans S. Treating problem-solving deficits on an acute care psychiatric inpatient unit. Psychiatry Research. 2000;97(1):79-88.
63. Wykes T, Reeder C, Huddy V, Taylor R, Wood H, Ghirasim N, et al. Developing models of how cognitive improvements change functioning: mediation, moderation and moderated mediation. Schizophr Res. 2012;138(1):88-93.
64. Cella M, Wykes T. The nuts and bolts of Cognitive Remediation: Exploring how different training components relate to cognitive and functional gains. Schizophr Res. 2017;203:12-16.
65. Crawford MJ, Robotham D, Thana L, Patterson S, Weaver T, Barber R, et al. Selecting outcome measures in mental health: the views of service users. Journal of Mental Health. 2011;20(4):336-46.
66. Rose D, Wykes TIL, Farrier D, Doran A-M, Sporle TIM, Bogner D. What Do Clients Think of Cognitive Remediation Therapy?: A Consumer-Led Investigation of Satisfaction and Side Effects. American Journal of Psychiatric Rehabilitation. 2008;11(2):181-204.
67. Turner-Stokes L. Goal attainment scaling (GAS) in rehabilitation: a practical guide. Clinical rehabilitation. 2009;23(4):362-70.

68. Gruber SA, Dahlgren MK, Sagar KA, Gonenc A, Norris L, Cohen BM, et al. Decreased Cingulate Cortex activation during cognitive control processing in bipolar disorder. J Affect Disord. 2017;213:86-95.
69. Hajek T, Alda M, Hajek E, Ivanoff J. Functional neuroanatomy of response inhibition in bipolar disorders--combined voxel based and cognitive performance meta-analysis. J Psychiatr Res. 2013;47(12):1955-66.
70. Penades R, Gonzalez-Rodriguez A, Catalan R, Segura B, Bernardo M, Junque C. Neuroimaging studies of cognitive remediation in schizophrenia: A systematic and critical review. World J Psychiatry. 2017;7(1):34-43.
71. Ramsay IS, MacDonald AW 3rd. Brain Correlates of Cognitive Remediation in Schizophrenia: Activation Likelihood Analysis Shows Preliminary Evidence of Neural Target Engagement. Schizophr Bull. 2015;41(6):1276-84.
72. Macoveanu J, Demant KM, Vinberg M, Siebner HR, Kessing LV, Miskowiak KW. Towards a biomarker model for cognitive improvement: No change in memory-related prefrontal engagement following a negative cognitive remediation trial in bipolar disorder. Journal of Psychopharmacology. 2018;32(10):1075-85.
73. Meusel LA, Hall GB, Fougere P, McKinnon MC, MacQueen GM. Neural correlates of cognitive remediation in patients with mood disorders. Psychiatry Res. 2013;214(2):142-52.

Tables

Table 1. Characteristics and findings of cognitive enhancing studies in bipolar disorder.

Study	Design	Sample characteristics	Clinical characteristics	Screening for impairment	Outcome measures	Main findings and limitations
Deckersbach et al., 2010 (28)	Uncontrolled study CR (N=18) BD: 100%	18 outpatients Age (y): 36.8 Gender (% fem): 55.5 Education (y): 14.5 Premorbid IQ: 105.9	BD-I (%): 82.3 Illness duration (y): Not reported Mood episodes (n): Not reported Hospitalizations (n): Not reported Current mood: Residual symptoms	None	Cognition: RBANS, D-KEFS subtests (TMT, CS), FrSBe Functioning: HPQ, LIFE-RIFT	An observer-based measure of executive dysfunction improved significantly post-treatment. Occupational and general psychosocial functioning also improved after 14 weeks of CR. Changes in executive functioning was not associated with functional improvement. Small sample size. Uncontrolled design.
Preiss et al., 2013 (24)	Non-randomized controlled trial CT vs. SC (N=24/21) BD: 46.6% MDD: 53.4%	45 outpatients Age (y): 42.9/45.4 Gender (% fem): 66.7/56.3 Education (y): Not reported Premorbid IQ:	BD-I (%): Not reported Illness duration (y): Not reported Mood episodes (n): Not reported Hospitalizations (n): Not reported	None	Cognition: CogniFit software tests Self-report cognition: CFQ, DEX, EMQ	CT was associated with improvement in mood and executive control in a mixed sample of depressed patients. A significant group effect was detected in subsets of Divided Attention (d=0.94), Shifting (d=0.88), and Global Executive Control score (d=0.9). Self-report cognitive measures (DEX/CFQ) showed a trend for the CT group.

		Not reported	Current mood: Depressed		Quality of life: SQUALA	Non-randomized design. Pairwise allocation per diagnosis. Small sample size. High attrition rates (24-37%). No FU assessment.
Torrent et al., 2013 (29)	RCT FR vs. PE / TAU (N=77/82/80) BD: 100%	239 outpatients Age (y): 40.6/39.3/40.5 Gender (% fem): 57.1/58.5/57.5 Education (y): 12.6/13.3/13.2 Premorbid IQ: 105.9/103.2/107.7	BD-I (%): 78/77/79 Illness duration (y): 14.8/12.7/16.4 Mood episodes (n): 11.7/9.9/13 Hospitalizations (n): 2.8/2.6/2.5 Current mood: Euthymic	Subjective criteria: Score ≥4 in FAST cognitive subscale	Cognition: VC, DSST, SS, SCWT, WCST, COWAT, TMT, ROCF, CVLT, LMS, DS, LNS, CPT Functioning: FAST	The FR group demonstrated greater improvement over TAU in psychosocial functioning (FAST; d=0.3), an effect driven by changes in interpersonal and occupational functioning. Only a trend was found for the FR group compared to PE. Although most cognitive measures improved numerically, no significant differences were found between groups at study endpoint. High trial attrition rate (17-29%). Unclear if ITT conducted. No FU assessment. Functional improvement not associated with cognitive gains.

Bonnin et al., 2016a (30)	Follow-up analysis of RCT FR vs. PE / TAU (N=54/60/58) BD: 100%	172 outpatients Age (y): 40.6/39.3/40.5 Gender (% fem): 57.1/58.5/57.5 Education (y): 12.6/13.3/13.2 Premorbid IQ: 105.9/103.2/107.7	BD-I (%):78/77/79 Illness duration (y): 14.8/12.7/16.4 Mood episodes (n): 11.7/9.9/13 Hospitalizations (n): 2.8/2.6/2.5 Current mood: Euthymic	Subjective criteria: Score ≥4 in FAST cognitive subscale	Cognition: VC, DSST, SS, SCWT, WCST, COWAT, TMT, ROCF, CVLT, LMS, DS, LNS, CPT Functioning: FAST	A small FR group effect on FAST was maintained over TAU and was extended to PE at 1-year FU (d=0.18), mainly due to reduction of difficulties in the autonomy domain. An improvement in verbal memory was detected for the FR group only (d=0.2), and it was associated with functional changes. High trial attrition rate (28-30%). No ITT analysis. Unblinded assessments.
Bonnin et al., 2016b (31)	Subanalysis of RCT FR vs. PE / TAU (N=56/69/63) BD: 100%	188 outpatients Age (y): 40.7/39.3/39.9 Gender (% fem): 58.9/60.9/58.7 Education (y): 12.3/12.8/12.8 Premorbid IQ: 102.8/101.7/104.8	BD-I (%): 77/73/77 Illness duration (y): 14.3/12.7/15.7 Mood episodes (n): 7/6.4/7.7 Hospitalizations (n): 3/2.6/2.6 Current mood: Euthymic	Objective criteria: Score ≥2 SDs from the normative mean in one cognitive test	Cognition: VC, DSST, SS, SCWT, WCST, COWAT, TMT, ROCF, CVLT, LMS, DS, LNS, CPT Functioning: FAST	Participants with objective cognitive impairment in the FR group significantly improved in verbal memory compared to TAU (d=0.3) but not PE. A trend for improvement in verbal learning was detected for the FR group only. Functional difficulties also reduced (d=0.2) but this change was not associated with cognitive improvement. No ITT analysis. No accounting for multiple comparisons. No FU assessment.
Sole et al., 2015	Subanalysis of RCT	53 outpatients Age (y):	BD-I (%):0/0/0 Illness duration (y):	Subjective criteria:	Cognition: VC, DSST, SS, SCWT,	In a subset of participants with bipolar type II, functional difficulties in the FR group reduced numerically but not

(32)	FR vs. PE / TAU (N=17/19/17) BD: 100%	42.6/40.2/40.4 Gender (% fem): 58.8/52.6/64.7 Education (y): 12/13.1/13.9 Premorbid IQ: 105.6/105.7/107.7	16.1/14.8/15.1 Mood episodes (n): 8.7/8.7/10.1 Hospitalizations (n): 3/2.8/2.2 Current mood: Euthymic	Score ≥4 in FAST cognitive subscale	WCST, COWAT, TMT, ROCF, CVLT, LMS, DS, LNS, CPT Functioning: FAST	significantly over PE or TAU. A significant reduction in depressive symptoms was detected compared to PE only. No significant effects were found for any of the neuropsychological measures. Small sample size. High trial attrition rates (12-29%). No ITT analysis. No FU assessment.
Sanchez-Moreno et al., 2017 (33)	Subanalysis of RCT FR vs. PE / TAU (N=33/37/29) BD: 100%	99 outpatients Age (y): 40.6/40.5/39 Gender (% fem): 45.5/48.6/48.3 Education (y): 13.3/12.6/13 Premorbid IQ: 105/100.2/105.7	BD-I (%): 61/75/62 Illness duration (y): 16/11.4/14.6 Mood episodes (n): 7.5/5.1/7.5 Hospitalizations (n): 2.6/2/2.8 Current mood: Residual symptoms	Subjective criteria: Score ≥4 in FAST cognitive subscale	Cognition: VC, DSST, SS, SCWT, WCST, COWAT, TMT, ROCF, CVLT, LMS, DS, LNS, CPT Functioning: FAST	A large group effect was found for FR in psychosocial functioning compared to PE and TAU in a subgroup of participants with residual depressive symptoms (d=1.12). No significant effects were reported for any of the cognitive measures, but a trend for improvement in verbal learning and problem solving was detected for the FR group. Attrition not reported. Unclear if ITT conducted. Functional improvement not associated with cognitive changes.
Demant et al., 2015 (26)	RCT CR vs. ST (N=23/23)	46 outpatients Age (y): 33.9/34 Gender (% fem): 66.7/59.1	BD-I (%): 72.2/63.6 Illness duration (y): Not reported Mood episodes (n):	Subjective criteria: Score >4 in ≥2 domains	Cognition: RAVLT, TMT, DSST, DS, LNS, COWAT, CANTAB (RVP, DMS, SWM,	No significant post-treatment group effects on verbal learning and memory, or any other measures of cognitive or psychosocial functioning, were detected in participants with subjective cognitive complaints. An improvement for the CR

	BD: 100%	Education (y): 15.9/15.7 Premorbid IQ: Not reported	4.9/9.2 Hospitalizations (n): Not reported Current mood: Euthymic	of the CPFQ	SRT), CFQ Functioning: FAST, CPFQ, WHOQOL, WSAS, EQ-5D-3L	group in verbal fluency (d=0.25) and psychological quality of life (WHOQOL; d=0.75) was reported at 6-month FU. Relatively small sample size. No ITT analysis. Low CR intensity. High attrition rate in the CR arm (21.7%).
Zyto et al., 2016 (25)	Uncontrolled study FR (N=12) BD: 100%	12 outpatients Age (y): 50.2 Gender (% fem): 54.5 Education (y): 4.6 in ISCED Premorbid IQ: 109.2	BD-I (%): 100 Illness duration (y): 28.6 Mood episodes (n): 5 Hospitalizations (n): 1.5 Current mood: Euthymic	None	Cognition: DS, DSST, TMT, SCWT, RAVLT, 15-WT, RBMT, WCST, COWAT, CFQ Functioning: FAST	Combined group and individual FR was feasible to deliver and acceptable from patients and caregivers. Psychosocial functioning improved, particularly the autonomy domain. Cognitive complaints significantly reduced at post-treatment, but the effect did not endure at FU. Small sample size. Uncontrolled design.
Lewandowski et al., 2017 (23)	RCT CR vs. CC (N=39/36) BD: 100%	75 outpatients Age (y): 29.3/29.8 Gender (% fem): 51/58 Education: 5.5/5.2 in SCID Premorbid IQ:	BD-I (%): 100/100 Illness duration (y): 7.5/8.5 Mood episodes (n): Not reported Hospitalizations (n): 4.8/3.8	None	Cognition: MCCB Functioning: MCAS	CR group was associated with a significant effect over computer control in multiple cognitive domains, including the composite score (d=0.8). Effect sizes were moderate to large (d=0.42-0.92) and remained significant at 6-month FU. These effects were independent to any symptom changes. Cognitive improvement was correlated significantly with changes in community functioning across the sample.

		113.4/112.1	Current mood: Euthymic			No ITT analysis. High CR discontinuation rates (33%). Correlation of cognitive and functional changes was not specific to CR group.
Veeh et al., 2017 (27)	Naturalistic study CR vs. TAU (N=26/10) BD: 100%	36 outpatients Age (y): 42.3/36.6 Gender (% fem): 50/50 Education (y): 11.8/11.6 Premorbid IQ: 111.5/113.3	BD-I (%): 56.3/50 Illness duration (y): Not reported Mood episodes (n): Not reported Hospitalizations (n): Not reported Current mood: Euthymic	Objective criteria: Score below the normative mean in two cognitive tests	Cognition: SCWT, TAP, CVLT, 2-back, ToL, FLEI Functioning: SFS, WHOQOL	A significant group effect was detected for CR in measures of problem solving (d=0.75) and working memory (d=0.66). An improvement was also reported for subsyndromal depressive symptoms. No group differences were found in measures of psychosocial functioning and subjective cognitive complaints. Non-randomized design. Self-selection of group allocation. Small sample size. Low treatment intensity. High attrition rate in the intervention arm. No FU assessment.

Abbreviations: *Design* (RCT: Randomized controlled trial; BD: Bipolar disorder; MDD: Major depressive disorder; FU: Follow-up; SD: Standard deviation; ITT: Intention-to-treat), *Interventions* (CR: Cognitive remediation; CT: Cognitive training; FR: Functional remediation; PE: Psychoeducation; CC: Computer control; TAU: Treatment-as-usual; SC: Standard care; ST: Standard treatment), *Education* (SCID: Structured Clinical Interview for DSM-IV; ISCED: International Standard Classification of Education), *Cognitive measures* (15-WT: 15 word test; COWAT: Controlled oral word association test; CPT: Continuous performance test; CVLT: California verbal learning test; CS: Card sorting; DMS: Delayed matching to sample; DS: Digit span; DSST: Digit-symbol substitution test; LMS: Logical memory scale; LNS: Letter number sequencing; MCCB: Matrics consensus cognitive battery; RAVLT: Rey-Auditory verbal learning test; RBANS: Repeatable battery of the assessment of neuropsychological status; RBMT: Rivermead behavioral memory test; ROCF: Rey-Osterrieth Complex Figure; RVP: Rapid visual information

processing; SCWT: Stroop color-word interference test; SRT: Simple reaction time; SS: Symbol search; SWM: Spatial working memory; TAP: Test of attentional performance; TMT: Trail making test; ToL: Tower of London; WCST: Wisconsin card sorting test; VC: Vocabulary test; ZMT: Zoo map test), *Self-report cognitive measures* (CFQ: Cognitive failures questionnaires; DEX: Dysexecutive questionnaire; EMQ: Everyday memory questionnaire; FLEI: Self-assessment test of mental ability; FrSBe: Frontal system behavior rating scale), *Functional measures* (CPFQ: Cognitive and physical functioning questionnaire; EQ-5D-3L: European quality of life – 5 dimensions – 3 levels; FAST: Functional assessment short test; HPQ: Health performance questionnaire; LIFE-RIFT: Longitudinal interval follow-up evaluation-Range of impaired functioning tool; MCAS: Multnomah community ability scale; SFS: Social functioning scale; SQUALA: Subjective quality of life questionnaires; WHOQOL: WHO Quality of life test; WSAS: Work and social adjustment scale).

Table 2. Description of the interventions targeting cognition and functioning in bipolar disorder.

Study	Intervention	Delivery method	Setting	Duration	Therapeutic targets	Core therapy components
Deckersbach et al., 2010	Compensatory cognitive remediation	Non-computerized training	Individual	50 min weekly for 3 months & biweekly for 1 month Total: 14 sessions	Residual symptoms Memory, Attention Organization, Planning	Training of cognitive skills with adaptive level of difficulty. Strategy learning focused on daily life management. Mood monitoring.
Preiss et al., 2013	Personalized restorative cognitive training	Computerized (CogniFit)	Individual	30 min three times per week for 8 weeks Total: 24 sessions	Multiple cognitive domains	Training personalized based on baseline evaluation. Cognitive tasks with adaptive level of difficulty. Graphic and verbal feedback.
Torrent et al., 2013 *(and all secondary analyses)*	Functional remediation	Pen-and-paper tasks and group activities	Group	90 min weekly for 21 weeks Total: 21 sessions	Daily life functioning, Memory, Attention, Executive functions	Psychoeducation on cognitive deficits and training on strategies to manage cognitive difficulties. Role-playing, group activities and discussions, and homework tasks to improve various aspects of daily-life

						functioning.
Demant et al., 2015	Compensatory cognitive remediation	Computerized (RehaCom)	Group	120 min weekly sessions for 3 months Total: 12 sessions	Attention-Concentration, Learning-Memory Executive functions, Psychosocial functioning	Psychoeducation and awareness of cognitive deficits. Computer practising and training of adaptive and compensatory strategies. Transfer of training to real-life activities. Mindfulness exercises.
Zyto et al., 2016	Functional remediation	Non-computerized individual training and group activities	Individual & group	90 min weekly for 6 weeks & 45 min weekly for 6 weeks Total: 12 sessions	Psychosocial functioning, Processing speed Memory, Attention, Planning	Personalized goal setting and strategy learning to cope with cognitive difficulties. Gaining insight and challenging dysfunctional thoughts. Group discussions.
Lewandowski et al., 2017	Neuroplasticity-informed cognitive remediation (restorative)	Computerized (BrainWorks)	Individual	45 min three times per week for 6 months Total: Up to 70 sessions	Global cognition and community functioning	Computer practising with games of adaptive difficulty level based on user performance. Bottom-up training: basic sensory processing followed by higher level functions. Rewarding for correct responses.
Veeh	Compensatory cognitive	Computerized	Individual	90 min weekly	Memory, Attention,	Computer tasks with adaptive

et al., 2017	remediation	(HappyNeuronPro)	& group	for 3 months Total: 12 sessions	Problem solving, Planning, Healthy lifestyle	difficulty level. Strategy training with connection to daily-life tasks. Positive reinforcement to maintain motivation levels. Graphic and verbal feedback.

Table 3a. Methodological evaluation of the included controlled studies using CTAM.

Study	Sample (0-10)	Allocation (0-16)	Assessment (0-32)	Control (0-16)	Analysis (0-15)	Treatment (0-11)	CTAM Total (0-100)
Preiss et al., 2013	2	0	13	6	5	3	**29**
Torrent et al., 2013	7	10	26	16	9	6	**74**
Demant et al., 2015	5	13	26	6	9	6	**62**
Sole et al., 2015	2	10	26	16	5	6	**65**
Bonnin et al., 2016a	7	10	16	16	5	6	**60**
Bonnin et al., 2016b	7	10	26	16	5	6	**70**

Lewandowski et al., 2017	7	13	26	10	9	3	**71**
Sanchez-Moreno et al., 2017	7	10	26	16	5	6	**70**
Veeh et al., 2017	2	0	26	6	5	3	**42**

Abbreviations: CTAM: Clinical Trials Assessment Measure.

Table 3b. Risk of bias for the included RCTs using the Cochrane Collaboration tool.

Study	Potential sources of bias							Global risk of bias judgement
	1	2	3	4	5	6	7	
Torrent et al., 2013	+	+	?	+	+	+	No (high drop-out rate)	Moderate
Demant et al., 2015	+	+	–	+	+	+	No (low CR retention rate; no ITT analysis)	High
Sole et al., 2015	+	+	?	+	–	+	No (small sample size; no ITT analysis)	Moderate
Bonnin et al., 2016a	+	+	–	?	–	+	No (high drop-out rate; no ITT analysis)	High
Bonnin et al., 2016b	+	+	?	+	–	+	No (no ITT analysis)	Moderate
Lewandowski et al., 2017	+	+	?	+	+	+	No (high drop-out rate)	Moderate
Sanchez-Moreno et al., 2017	+	+	?	+	–	+	No (no ITT analysis; drop-out rate not reported)	Moderate

1. Selection bias: Random sequence generation; 2. Selection bias: Allocation concealment; 3. Performance bias: Blinding of participants & personnel; 4. Detection bias: Blinding of outcome assessment; 5. Attrition bias: Adequate handling of missing data; 6. Reporting bias: Free of selective outcome reporting; 7. Other bias: Free of other sources of bias. +, Low risk / –, High risk / ?, Unclear risk.

Abbreviations: BD: Bipolar disorder; ITT: Intention-to-treat analysis.

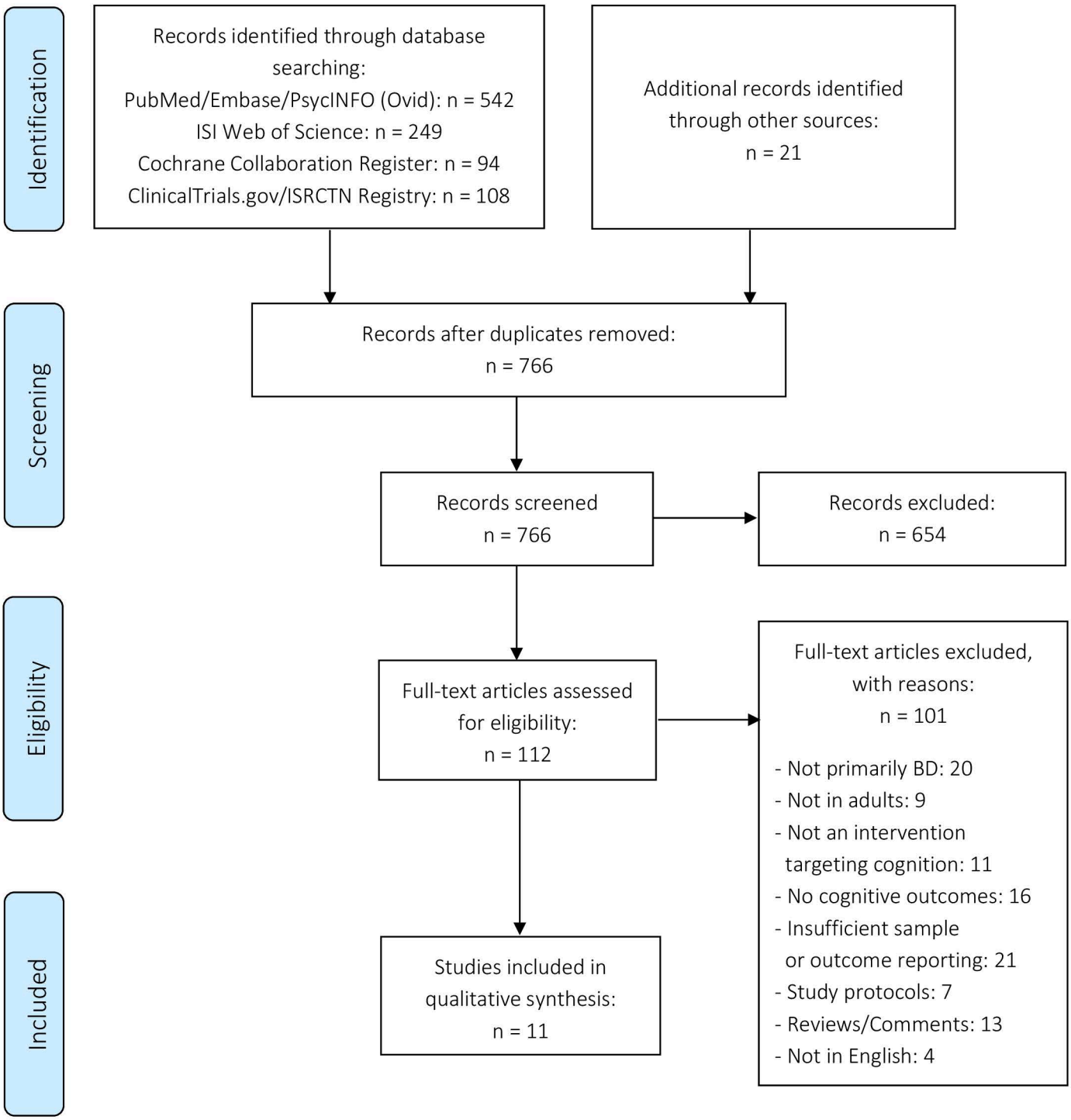
Table 4. Categorising studies by design components potentially contributing to outcome variability.

Design components	Factors	Factor levels	Number of studies
Sample characteristics	Diagnosis	BD Type I & Type II	8
		BD Type I only	2
		BD Type II only	1
	Mood symptoms at baseline	Euthymic	8
		Residual depressive symptoms	2
		Depressed	1
	Baseline cognitive screening	No screening	4
		Subjective screening	5
		Objective screening	2
Assessments	Cognitive measures	Standardized battery	2
		Individual tests	9
	Composite score	Reported	1
		Not reported	10
	Treatment feedback	Examined	2
		Not examined	9
Intervention	Approach	Compensatory	9
		Restorative	2
	Setting	Group	6
		Individual	3
		Combination	2
	Delivery	Pen & paper	7
		Computer	4
	Intensity/duration	Low/long	5
		Low/medium-short	4
		High/short	1
		High/long	1
	Therapist	Yes	9
		No	2

Abbreviations: BD: Bipolar disorder.

Figure legends

Figure 1. PRISMA flow diagram presenting study identification and selection


Identification
Screening
Eligibility
Included
Records identified through database searching:
PubMed/Embase/PsycINFO (Ovid): n = 542
ISI Web of Science: n = 249
Cochrane Collaboration Register: n = 94
ClinicalTrials.gov/ISRCTN Registry: n = 108
Additional records identified through other sources:
n = 21
Records after duplicates removed:
n = 766
Records screened
n = 766
Records excluded:
n = 654
Full-text articles assessed for eligibility:
n = 112
Full-text articles excluded, with reasons:
n = 101
- Not primarily BD: 20
- Not in adults: 9
- Not an intervention targeting cognition: 11
- No cognitive outcomes: 16
- Insufficient sample or outcome reporting: 21
- Study protocols: 7
- Reviews/Comments: 13
- Not in English: 4
Studies included in qualitative synthesis:
n = 11