Treatment to Enhance Cognition in Bipolar Disorder (TREC-BD): Efficacy of a Randomized Controlled Trial of Cognitive Remediation versus Active Control

Kathryn E. Lewandowski, PhD1,2, Sarah H. Sperry, MA1, Bruce M. Cohen, MD, PhD1,2, Lesley A. Norris, BA1, Garrett M. Fitzmaurice, PhD2,3,4, Dost Ongur, MD, PhD1,2 &,

Matcheri S. Keshavan, MD2,5

1 Schizophrenia and Bipolar Disorder Program, McLean Hospital, Belmont, MA, USA
2 Department of Psychiatry, Harvard Medical School, Boston, MA, USA
3 Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA
4 Psychiatric Biostatistics Laboratory, McLean Hospital, Belmont, MA, USA
5 Department of Psychiatry, Beth Israel Deaconess Medical Center, Boston, MA, USA

This work was supported by the National Institute of Mental Health (K23MH091210 to KEL; K24MH104449 to DO).

Address Correspondence To:
Kathryn E. Lewandowski
McLean Hospital/Harvard Medical School
115 Mill St. AB South 349, Belmont, MA 02478
klewandowski@mclean.harvard.edu

Abstract
Objective: Cognitive dysfunction is a core symptom dimension in bipolar disorder, and a strong predictor of functional outcomes. Cognitive remediation (CR) produces moderate, durable effects on cognition in patients with schizophrenia; however, studies of CR in patients with bipolar disorder are sparse and findings have been mixed. Thus, we aimed to evaluate the effects of CR versus active control in patients with bipolar disorder with psychosis. Methods: Patients with a DSM-IV diagnosis of bipolar disorder with psychosis (n=75) were randomized to a 70-hour computerized CR program or a dose-matched computer control using a parallel design with 1:1 allocation between July, 2011 and November, 2015. Cognition (primary) and clinical and community functioning (secondary) were assessed at baseline, treatment midpoint (20-25 hours), post-treatment, and durability (after 6 months of no study contact). Participants and assessment staff were blind to group membership. Results: 75 participants were randomized and 72 participants initiated the active phase of treatment and were included in the primary, intent-to-treat analysis (CR: n=39; Control: n=33). Linear mixed effects models examining the effects of CR versus control at post-treatment showed medium to large effects of CR on processing speed (d=.42), visual learning and memory (d=.92), and the composite (d=.80). Superiority of CR over control on processing speed (d=.65) and composite (d=.83) were maintained or increased at durability. CR was not associated with change in community functioning, although cognitive change was associated with functional change across the sample. Conclusion: CR produced significant improvements over control in several cognitive domains and the cognitive composite. While both groups improved on several domains relative to baseline, durability of gains was unique to CR. Trial Registration: ClinicalTrials.gov identifier: NCT01470781
INTRODUCTION

Cognitive dysfunction is a core symptom dimension in bipolar disorder (BD), and is a strong predictor of disability, poor quality of life, and longer time to recovery after a first episode. As few as one third of patients achieve functional recovery over time. At least partial disability is reported in approximately 80% of patients with BD; as many as 65% of patients are unemployed even after clinical recovery, and continue to experience significant disability in daily living and social functioning. Given the association between cognition and functional outcomes, the cognitive symptom dimension represents a critical treatment target in BD.

Cognitive Deficits in Bipolar Disorder

Patients with BD exhibit deficits in memory, executive function, and processing speed that persist during euthymic phases and over time. At the group level, cognitive deficits in patients with BD are qualitatively similar to those of patients with schizophrenia (SZ) and related disorders, and some findings report similar magnitudes of dysfunction, particularly in patients with a history of psychosis. Recent cluster analyses report similar cognitive subgroups in BD and SZ, including a globally impaired cluster characterized by severe cognitive impairment across domains.

Treating Cognitive Symptoms

Cognitive remediation (CR) is a behavioral intervention targeting the cognitive symptom dimension to improve cognitive functioning. On average, CR paradigms produce moderate, durable effects on cognitive performance in patients with SZ and related illnesses. A recent review reported moderate effects (0.32) of CR on cognition in affective illness, primarily affective psychosis.
Despite considerable overlap in cognitive symptoms between patients with SZ and BD, reports of CR outcomes in patients with BD are sparse. An open trial of a 14-session CR program for patients with BDI or BDII\textsuperscript{18} was associated with improved executive functioning and vocational performance. However, a study by Demant and colleagues\textsuperscript{19} failed to show an effect of a 12-week group-based CR program compared to treatment as usual in patients with BD. A functional remediation program\textsuperscript{20} improved verbal memory compared to treatment as usual.\textsuperscript{21} This program differed from neuroplasticity-based CR in several ways, including group-based training and an emphasis on role play and instruction. Thus, CR findings are mixed in patients with BD, and the heterogeneity of study designs limits both the interpretability of the efficacy of CR in BD and the comparability of outcomes to other psychiatric populations.

Inclusion of adequate comparison groups is a particular challenge in CR, as participants and study administrators are aware of the elements of participation by nature of the intervention. Active controls most often involve “open-label” rehabilitative or therapeutic activities.\textsuperscript{22} While such controls permit evaluation of the effects of some non-specific elements of the intervention (e.g. hours of study contact), they cannot control for other potentially impactful elements of CR (e.g. engagement with computer-based activities; expectancies). Lack of double-blindness presents challenges in interpreting results,\textsuperscript{23} as poor masking has been shown to inflate treatment effects.\textsuperscript{16} To our knowledge, there are no reported studies attempting to test a “blind” comparison condition in patients with BD.

The present study aimed to evaluate the effects on cognition of a randomized, double-blind controlled trial of computer-based CR in patients with BDI with a history of
psychosis (BDP) compared to a computer control. We chose a CR program (PositScience) that has been shown to be effective in patients with SZ.\textsuperscript{24} We hypothesized CR would produce cognitive improvement compared to active control. Our secondary aims were to examine associations between CR and clinical and community outcomes. It was hypothesized that change in cognition but not in clinical state would predict improvements in community functioning. Lastly, we aimed to evaluate whether our active control effectively maintained the blind.

**METHOD**

This study and all associated procedures comply with the Helsinki Declaration of 1975, as revised in 2008, and were approved by the McLean Hospital IRB. Details of the study protocol have been described previously.\textsuperscript{22} This project is registered with ClinicalTrials.gov (NCT01470781), and all procedures and reporting follow CONSORT 2010 guidelines.\textsuperscript{25}

*Participants*

Participants ages 18-50 with BDP (*n*=84) were enrolled through the McLean Hospital Schizophrenia and Bipolar Disorder Program and via fliers posted at the hospital between July 15, 2011 and November 18, 2015. Of the 84 participants enrolled, 75 were randomized and 72 initiated the active phase of the study (Figure 1). All participants had a diagnosis of BDP and were stable outpatients at the time of enrollment. We chose to include only patients with a history of psychosis for two primary reasons: 1) to reduce heterogeneity of our sample, given the considerable variability within BD generally, and 2) because some literature suggests that a history of psychosis in patients with BD is associated with more severe cognitive deficits. Medication exclusions included clozapine,
anticholinergic medications, and topiramate due to potential effects on CR response.\textsuperscript{26} Other exclusion criteria included DSM-IV-TR substance abuse in the past month or dependence in the past year, history of seizure disorder, or history of head injury with a loss of consciousness. All subjects provided written informed consent after receiving a complete description of the study.

\textit{Materials}

Cognitive functioning was assessed using the MATRICS Consensus Cognitive Battery (MCCB\textsuperscript{27}), which has been validated for use in patients with BD.\textsuperscript{10, 28} The MCCB includes 10 tasks that measure processing speed, attention, working memory, verbal learning, visual learning, problem solving, and social cognition. Total administration time was 60-90 minutes.

Clinical and community functioning assessment included the PANSS\textsuperscript{29}, YMRS\textsuperscript{30}, Montgomery-Asberg Depression Rating Scale (MADRS\textsuperscript{31}), and an abbreviated version\textsuperscript{11} of the Multnomah Community Ability Scale (MCAS\textsuperscript{32}), which measures functioning in multiple domains (e.g. social, independence, role functioning). A home-grown 13-item user feedback survey measured participants’ experience of the CR or Control activities; cumulative scores range from 13-65. Participants were then asked to indicate their belief about their study assignment (binary forced choice), and to rate the certainty of their selection.

\textit{Procedures}

\textit{Randomization and Masking.} Participants were randomly assigned to CR or Control using a parallel (1:1) blocked randomized design, block size of ten. Randomization assignments were generated using a computerized randomization generator; using this pre-specified structure, treatment administrators were responsible for assignment. Due to
the nature of the intervention, treatment administrators were not able to remain blind to group assignment; however, assessment staff, investigators, and participants were blind to group assignment. Assessment staff and investigators were never involved in the randomization or treatment administration procedures. The blind was maintained until the final participant completed the durability assessment.

CR intervention. The CR protocol involved the BrainWorks programs by PositScience. Games were developed based on a recovery model of neural plasticity, using a “bottom up” approach to train sensory processing during the early weeks of training and adding "higher-order" tasks as the program progressed. Programs self-adjust based on user performance to keep participants working at 80% proficiency. Games include basic auditory and visual perception activities, tasks of divided attention, memory and working memory games, and problem solving games (Supplemental Table 1). Activities are packaged in a game format, and participants earn points and virtual rewards for correct responses.

Computer Control. We developed an internet-based control intervention designed to mirror the CR program in number of training sessions, administrator contact, and format to control for as many nonspecific effects of the CR treatment as possible without involving tasks designed to target the specific cognitive deficits addressed in the CR condition, as previously described. Thus, while participants were engaged in computer games actively, they were not undergoing a structured program designed to strengthen both basic sensory processing and higher order cognitive skills in a systematized way. Briefly, sessions were created and administered via the online interface Sporcle (www.Sporcle.com), a collection of quiz-type activities. Several hundred games were selected to include a broad array of
activities including identification activities (e.g. identify pictured fruits; name popular logos), content-based activities (e.g. label maps; pop-culture quizzes), and timed activities (e.g. basic arithmetic). Games were divided into 70 sessions, with activities distributed so that no single session included only one type of game. Participants were given “prescription cards” weekly assigning games by session.

Both conditions involved approximately 3 sessions per week over 24 weeks, with a target of 70 sessions. Participants attended one session per week at the study site; two sessions were completed remotely via internet login. Weekly in-person sessions included a CR or control game session plus a clinical update (CR and control) and a brief bridging session (CR only), which involved participant and staff discussions regarding participants’ application of skills in daily life since the previous session. Administration staff monitored CR and control participant activities remotely via an internet-based administrator account. Reminder calls were offered as needed. Participants earned $5 per completed session.

Assessments were conducted at baseline, midpoint (after 20-25 hours of training), post-treatment, and at durability (after 6 months of no study contact). All assessments involved completion of cognitive, clinical, and community functioning measures.

Statistical Approach

Groups were compared on baseline demographic, cognitive, and clinical variables using t-tests or $X^2$-tests. Linear mixed effects models predicting cognitive outcomes by group, time, and the group by time interaction were used in this intent-to-treat analysis. The control group and baseline assessment were coded as reference variables for group and time, respectively. Effect sizes of the mean change over time of the treatment group versus control using Cohen’s $d$ were calculated as $d = (\text{Mean Change}_{\text{treatment}} - \text{Mean Change}_{\text{control}})$.
In a supplementary analysis, number of completed sessions was included as a covariate to examine the effects of training hours.

Similar mixed effects models examined the effects of CR on community functioning based on group, time and the group by time interaction. Linear regressions predicting change in MCAS scores from baseline to post-treatment as a function of group, change in cognitive domain (baseline to post-treatment), and group by cognitive change interaction were conducted to examine the effects of cognitive change on community outcomes. Similar regressions were conducted using clinical change (in place of cognitive change) to examine effects of clinical change (post-baseline minus baseline) on community function.

Lastly, groups were compared on responses to the user feedback survey using t-tests or $X^2$ to evaluate tolerability of the activities and maintenance of participant blind.

**RESULTS**

CR and Control groups did not differ on any baseline characteristics, cognitive scores, chlorpromazine (CPZ) equivalents or lithium dose (Table 1), total number of medications prescribed at baseline, or proportion of patients on any given class of medication. CR and control participants were taking an average of 2.7 and 3.1 psychiatric medications, respectively. In the total sample 93% were taking a mood stabilizer, 69% were taking Lithium, 70% were taking an antipsychotic, 34% were taking an antidepressant, and 23% were taking a benzodiazepine.

Total number of sessions completed did not differ by group (CR: mean(SD)=43(26), range=4-70; Control: mean(SD)=48(25), range=4-70; $t=-0.81$, $p=0.42$). Six participants in the CR group and three in the control group completed fewer than 10 sessions.

*Effects of CR on Cognitive Outcomes*
Linear mixed effects models revealed significant group by time interactions at post-treatment for the cognitive composite, visual learning and memory, and a trend for processing speed indicating significant improvements of CR over control (Table 2; Supplemental Figure 1). At durability we found a significant effect of CR over control for the composite, processing speed, and trend for verbal learning and memory. Effect sizes were in the medium to large range for significant effects and in the medium range for trend level associations. Including the number of completed sessions, state mania, depression, or psychosis symptoms, CPZ, or Lithium equivalents as covariates did not change any of the findings.

Cognitive change scores from baseline to post-treatment and durability are presented in Figure 2. All significant findings were in the direction of greater improvements for CR compared to Control. In the CR group, standardized change scores showed mean improvements after CR at post-treatment of 0.59 SD in overall cognition, and 0.50 SD or greater for processing speed, attention, visual learning and memory, and problem solving. We found additional improvements at durability in the CR group in processing speed, working memory, verbal learning and memory, and the composite. The control group also showed improvements on several cognitive domains, although of smaller magnitude. At post-treatment, the composite improved 0.22 SD, and improvements in processing speed, attention, working memory, and problem solving were in the 0.3-0.4 SD range. The control group showed decline from post-training to durability on most measures.

**Effects of CR on Clinical Symptoms and Community Functioning.** Mixed effects models showed no effects of CR on any clinical measure or on MCAS scores. Linear regressions
predicting change in MCAS from baseline to post-treatment by group, cognitive change, and the group by cognitive change interaction showed a significant effect of cognitive change on community functioning in visual learning and memory (B=0.28, t=3.08, p=.004) and problem solving (B=0.19, t=2.63, p=.01) across the sample. Group by cognition interactions were not significant.

*Tolerability and Maintenance of the Blind*

User feedback data are reported in Supplemental Table 2. Mean scores of “positively-worded” items indicated moderate to high levels of satisfaction, and response to “negatively-worded” items were in the “rarely” to “never” range, suggesting good acceptability and tolerability of both conditions. Groups did not differ on any item, the total score, or beliefs about group assignment (χ² = 0.33; p=.85). Approximately half of the subjects in each group believed they were in the active treatment condition (CR: 10/19; Control: 5/10); groups reported equally low confidence in their ratings (CR = 1.7 (0.7); Control =1.5 (0.8); t=0.64; p=.53).

**DISCUSSION**

In this randomized, double-blind trial of CR in patients with BDP, CR produced significant effects over control in several cognitive domains including processing speed, visual learning and memory, and the cognitive composite, and a trend for verbal learning and memory at the durability assessment, with effects in the medium to large range. Note that these effects were measured against an active, dose-matched, computer-based control condition in order to rigorously test the effects of the CR program against other computer activities. This is the first study we are aware of to extend a neuroplasticity-informed CR paradigm previously shown to be effective in patients with SZ to patients with BDP and to
demonstrate its efficacy, and the first to employ a double-blind design in a CR paradigm in patients with BDP.

Effects were not likely due to changes in clinical status or medication over the course of the treatment. Change in YMRS, MADRS, and PANSS scores did not predict change in cognitive functioning in the total sample or by group. Total average medications did not differ by group and were essentially unchanged over the course of the study. However, in a post-hoc analysis we examined the effects of premorbid IQ on treatment response by including NAART scores in the mixed effects models and found that, while inclusion of premorbid IQ did not substantively change our primary findings, NAART score was positively associated with treatment response in the model ($z=2.04$, $p=.04$), suggesting that cognitive reserve may be an important moderator of response to CR interventions. Future studies should include measures of cognitive reserve and evaluate its impact on treatment response.

The effect on visual learning and memory at post-treatment was stronger than may have been expected based on previous work using this CR program. The version of the CR paradigm we used included activities that explicitly train basic and complex visual processing training (see Supplemental Material), interleaved with the auditory training which may have generated additional benefits in visual learning and memory. Alternatively, it is possible that patients with BD show greater gains in visual memory than patients with SZ. Future work is needed to determine whether effects are due to the exercises, or if there are true diagnostic differences in treatment response in visual domains.

*Durability of CR in BDP*
While CR produced robust effects on cognition compared to control, the control condition also resulted in improved cognition in several domains. However, only the CR group maintained this higher level of cognitive performance at durability. Most scores and the composite actually continued to improve during the durability period in the CR group, whereas most domain scores and the composite declined from post-treatment to durability in the control group. These findings are consistent with a one-year follow up of functional remediation,21,34 in which CR was equal to group psychoeducation at post-treatment, but at follow-up gains persisted in the CR group only. The mechanism(s) of action of CR and control may differ, even when both groups demonstrate short-term cognitive improvement. CR training may set participants on a trajectory for continued gains, which may be due to improved neurobiological functioning, increased engagement in cognitively stimulating activities, or both. Conversely, repetitive engagement in the control activities may have temporarily improved performance, but after cessation of the activities the gains were not maintained. Exploration of the neurobiological mechanisms of action will help to clarify how various “ingredients” act on specific biological and behavioral targets.

Specific Versus Non-specific Effects: CR versus Placebo

Our findings that the control group exhibited improvements – albeit to a lesser extent than CR – in several cognitive domains, suggests that nonspecific training elements may also improve cognitive performance. A recent study23 showed effects of participant expectancies on cognition after a brief cognitive training based on whether the activities were advertised as effective at enhancing cognition or not. These findings suggest that expectancies – perhaps via activation of reward systems35,36 – may drive improvements in cognitive performance, at least in the short-term. Given findings of placebo effects in CR
and the potential for poor masking to inflate treatment response,\textsuperscript{16} inclusion of blinded controls is key to our understanding of the efficacy and mechanisms of effect of CR.

\textit{Dosing}

Optimized dosing is essential to any intervention. In CR, “side effects” of treatment such as fatigue and partial- or non-adherence likely increase with dose. Indeed, the top reason given for early discontinuation was the time requirement. However, midpoint assessments indicated that 20-25 hours of training were not enough to demonstrate an effect of CR over control, suggesting that longer duration of CR is needed. Optimization of CR dosing is a critical next step, with consideration of both CR paradigm and patient characteristics. For instance, while multi-domain training may not be broadly effective at lower doses, domain specific training may drive change after fewer training hours.\textsuperscript{37}

\textit{CR and Community Functioning}

Community functioning was associated with change in cognition over time across the sample; however, these effects were not specific to CR. This suggests that improvements in community functioning track with improved cognition, regardless of the mechanism of the cognitive change. Our findings are consistent with reports\textsuperscript{38–40} that CR paradigms that include explicit rehabilitation elements show better transfer to functional outcomes than programs that offer drill-and-practice only. Alternatively, our measure of community functioning may not adequately tap relevant domains likely to change with improved cognition after CR. More detailed measures will help answer key questions regarding which domains of cognitive improvement are associated with particular community outcomes.

\textit{Limitations}
There are several limitations of the present study; most notably, the non-completion rate was high in both groups, reducing power to detect group effects. Thus, findings from this study should be considered preliminary. Nevertheless, we found significant effects of CR on several cognitive domains, and effect sizes comparable to or greater than those reported in other studies of CR in both affective illness and primary psychosis. However, high attrition rates also suggest that this intervention may not be well tolerated by a subset of participants. Reasons for non-completion included practical considerations (transportation/logistical issues), clinical factors (symptom exacerbations), and factors related to the treatment itself (games were frustrating, burdensome). Thus, many critical questions must be answered if we are to create programs that balance tolerability/feasibility with efficacy, including establishing clear dosing guidelines and identifying baseline or in-study factors associated with early discontinuation.

Additionally, in our sample, baseline cognitive scores and premorbid IQ were higher than would be expected based on the literature. Thus, our findings may not generalize to patients with cognitive functioning more typical of this population. However, a meta-analysis\textsuperscript{16} of CR in patients with SZ found no association between baseline cognition and cognitive outcomes. Also, our participants were young with an average duration of illness of eight years, which may limit the generalizability of our findings. Some findings in SZ\textsuperscript{41} suggest that younger patients may benefit more from CR interventions, although a meta-analysis\textsuperscript{16} found no effect of age or duration of illness on treatment response. Further work is needed to evaluate the effects of patient characteristics on CR response. Lastly, in keeping with a number of studies in patients with SZ, we failed to find strong evidence of transfer of gains to community functioning. Inclusion of supportive elements such as
rehabilitation has been associated with generalization of cognitive gains to functional changes;\textsuperscript{16,38,43} while we did include brief bridging sessions in our CR condition, it is possible that these were not adequate to drive translation of cognitive improvements to broad functional measures. An issue for consideration is the effect of incentives on participation. Our participants were paid a modest amount of money for completed sessions, which may have impacted treatment engagement. The role of motivation and incentives may be of particular importance to implementation in clinical settings.

Our findings indicate that CR is beneficial to patients with BDP and produces significant, durable cognitive change against an active, dose-matched computer control, supporting the extension of this intervention to patients with BD to address serious and disabling cognitive symptoms. Additionally, these findings support the implementation of web-based CR, as such treatments appear to be efficacious and offer therapeutic options that are cost effective and afford greater access than traditional clinic-based treatments \textsuperscript{42}. Further research is required to maximize transfer effects, and to match these cognitive effects to work and leisure skills and other aspects of community functioning in patients with BD.
Acknowledgments and Disclosures

This research was supported by the National Institute of Mental Health (K23MH091210 to KEL; R01MH092440 and R01MH78113 to MSK; K24MH104449 to DO). PositScience provided the CR training software for research purposes free of charge. Neither the funder of this study nor PositScience had any role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

KEL, DO, GMF, SHS, and LAN, declare that they have no financial relationships with commercial interests. BMC holds several patents, including: US patent 6,617,169; US patent 7,629,475; US patent 7,884,077; US patent 8,168,169; US patent 8,492,564, none of which are related to the present work. MSK was a one-time consultant for Forum Pharmaceuticals.

The authors wish to thank all of the participants for contributing their time and efforts to this project.
Clinical Points

- Cognitive dysfunction is a core symptom in bipolar disorder and a strong predictor of disability and poor quality of life.

- Cognitive remediation (CR) improves cognition in patients with schizophrenia; however, despite considerable overlap in cognitive symptoms between these disorders, there are few reports of CR outcomes in bipolar disorder.

- CR produced significant, lasting cognitive change in patients with bipolar disorder, supporting the extension of CR to this population.
REFERENCES


effects are defined by opposite opioid and dopaminergic responses. *Arch Gen Psychiatry*. 2008;65(2):220-231.


### TABLE 1. Baseline Demographic, Clinical, and Cognitive Characteristics by Group

<table>
<thead>
<tr>
<th>Variable</th>
<th>Treatment (n=39)</th>
<th>Control (n=33)</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>29.3 (7.5)</td>
<td>29.8 (9.2)</td>
<td>(t_{(70)}=0.23)</td>
</tr>
<tr>
<td>Gender (% female)</td>
<td>51%</td>
<td>58%</td>
<td>Chi^2(_{(1)}) = 0.29</td>
</tr>
<tr>
<td>Race (% Caucasian)</td>
<td>77%</td>
<td>85%</td>
<td>Chi^2(_{(1)}) = 0.72</td>
</tr>
<tr>
<td>Education(^a)</td>
<td>5.5 (1.6)</td>
<td>5.2 (1.4)</td>
<td>(t_{(70)}=0.89)</td>
</tr>
<tr>
<td>Duration of Illness</td>
<td>7.5 (6.5)</td>
<td>8.5 (8.1)</td>
<td>(t_{(68)}=-0.59)</td>
</tr>
<tr>
<td>Prior Hospitalizations</td>
<td>4.8 (5.3)</td>
<td>3.8 (4.0)</td>
<td>(t_{(70)}=0.92)</td>
</tr>
<tr>
<td><strong>PANSS Total</strong></td>
<td>47.5 (8.5)</td>
<td>45.4 (10.0)</td>
<td>(t_{(70)}=0.95)</td>
</tr>
<tr>
<td><strong>PANSS Positive</strong></td>
<td>10.5 (3.4)</td>
<td>10.5 (3.9)</td>
<td>(t_{(70)}=0.01)</td>
</tr>
<tr>
<td><strong>PANSS Negative</strong></td>
<td>11.6 (4.3)</td>
<td>10.4 (2.5)</td>
<td>(t_{(70)}=1.47)</td>
</tr>
<tr>
<td><strong>PANSS General</strong></td>
<td>25.4 (4.9)</td>
<td>24.5 (5.6)</td>
<td>(t_{(70)}=0.66)</td>
</tr>
<tr>
<td><strong>YMRS</strong></td>
<td>5.6 (4.9)</td>
<td>4.7 (4.5)</td>
<td>(t_{(70)}=0.79)</td>
</tr>
<tr>
<td><strong>MADRS</strong></td>
<td>11.8 (7.5)</td>
<td>12.2 (7.2)</td>
<td>(t_{(70)}=-0.20)</td>
</tr>
<tr>
<td><strong>CPZ</strong></td>
<td>183.2 (230.5)</td>
<td>151.4 (174.4)</td>
<td>(t_{(70)}=0.65)</td>
</tr>
<tr>
<td><strong>Li Dose</strong></td>
<td>669.2 (542.9)</td>
<td>900.0 (657.1)</td>
<td>(t_{(70)}=-1.63)</td>
</tr>
<tr>
<td><strong>Number of Meds</strong></td>
<td>2.79 (1.22)</td>
<td>3.06 (1.19)</td>
<td>(t_{(70)}=-0.94)</td>
</tr>
<tr>
<td><strong>MCAS</strong></td>
<td>47.7 (4.3)</td>
<td>48.3 (3.7)</td>
<td>(t_{(70)}=-0.67)</td>
</tr>
<tr>
<td><strong>NAART VIQ</strong></td>
<td>113.4 (8.0)</td>
<td>112.1 (8.0)</td>
<td>(t_{(56)}=0.62)</td>
</tr>
<tr>
<td><strong>MCCB Scores</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Processing Speed</td>
<td>46.6 (10.5)</td>
<td>47.9 (10.7)</td>
<td>(t_{(70)}=-0.52)</td>
</tr>
<tr>
<td>Attention</td>
<td>45.1 (12.5)</td>
<td>44.2 (10.5)</td>
<td>(t_{(70)}=0.29)</td>
</tr>
<tr>
<td>Working Memory</td>
<td>48.9 (10.7)</td>
<td>47.7 (8.0)</td>
<td>(t_{(70)}=0.56)</td>
</tr>
<tr>
<td>Verbal Learning</td>
<td>47.0 (10.9)</td>
<td>48.9 (7.8)</td>
<td>(t_{(70)}=-0.84)</td>
</tr>
<tr>
<td>Visual Learning</td>
<td>44.2 (10.2)</td>
<td>44.7 (11.4)</td>
<td>(t_{(70)}=-0.21)</td>
</tr>
<tr>
<td>Problem Solving</td>
<td>46.8 (9.4)</td>
<td>47.0 (10.6)</td>
<td>(t_{(70)}=-0.05)</td>
</tr>
<tr>
<td>Social</td>
<td>50.5 (8.9)</td>
<td>49.2 (10.9)</td>
<td>(t_{(70)}=0.54)</td>
</tr>
<tr>
<td>Composite</td>
<td>45.2 (10.4)</td>
<td>45.5 (9.12)</td>
<td>(t_{(70)}=-0.14)</td>
</tr>
</tbody>
</table>

PANSS: Positive and Negative Syndrome Scale; YMRS: Young Mania Rating Scale; MADRS: Montgomery-Asberg Depression Rating Scale; CPZ: Chlorpromazine Equivalent; Li: Lithium; MCAS: Multnomah Community Ability Scale; NAART: North American Adult Reading Test; VIQ: Verbal IQ; MCCB: MATRICS Consensus Cognitive Battery

\(^a\) Education is coded based on the SCID Education and Work History scale: 1=grade 6 or less; 2= grade 7-12 (without graduating); 3= high school grad or equivalent; 4= part college; 5= graduated 2 year college; 6= graduated 4 year college; 7= part graduate/professional school; 8= completed graduate/professional school.
<table>
<thead>
<tr>
<th>MCCB Domain</th>
<th>Midpoint Post-Treatment</th>
<th>95% CI</th>
<th>Effect Size</th>
<th>Durability</th>
<th>95% CI</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Processing Speed</strong></td>
<td>B= -0.37 (1.97)</td>
<td>-0.32 to 7.16</td>
<td>d= 0.42</td>
<td>B= 5.57 (2.19)</td>
<td>1.27 to 9.87</td>
<td>d= 0.65</td>
</tr>
<tr>
<td></td>
<td>z=0.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Attention</strong></td>
<td>B= -0.33 (2.53)</td>
<td>-6.79 to 5.48</td>
<td>d= 0.03</td>
<td>B= 2.64 (3.56)</td>
<td>-9.62 to 4.34</td>
<td>d= -0.25</td>
</tr>
<tr>
<td></td>
<td>z= -0.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>WMemory</strong></td>
<td>B= -1.26 (1.51)</td>
<td>-4.80 to 1.89</td>
<td>d= -0.07</td>
<td>B= 2.91 (1.99)</td>
<td>-0.98 to 6.81</td>
<td>d= 0.67</td>
</tr>
<tr>
<td></td>
<td>z= -0.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Verbal</strong></td>
<td>B= -1.77 (2.17)</td>
<td>-5.06 to 5.04</td>
<td>d= -0.05</td>
<td>B= 4.78 (2.77)</td>
<td>-0.65 to 10.22</td>
<td>d= 0.64</td>
</tr>
<tr>
<td></td>
<td>z= -0.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Visual</strong></td>
<td>B= -0.82 (2.49)</td>
<td></td>
<td></td>
<td></td>
<td>B= 2.87 (2.92)</td>
<td>-2.76 to 8.60</td>
</tr>
<tr>
<td></td>
<td>z= -0.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ProbSolv</strong></td>
<td>B= .889 (2.42)</td>
<td></td>
<td></td>
<td></td>
<td>B= 2.40 (2.83)</td>
<td>-3.15 to 7.94</td>
</tr>
<tr>
<td></td>
<td>z= 0.37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Social</strong></td>
<td>B= 1.12 (2.21)</td>
<td></td>
<td></td>
<td></td>
<td>B= 0.68 (2.56)</td>
<td>-5.70 to 4.33</td>
</tr>
<tr>
<td></td>
<td>z= 0.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>COMPOSITE</strong></td>
<td>B= -0.35 (1.23)</td>
<td></td>
<td></td>
<td></td>
<td>B= 5.07 (2.06)</td>
<td>1.03 to 9.12</td>
</tr>
<tr>
<td></td>
<td>z= -0.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

aResults of linear mixed effects models; Coefficient (standard error) and z statistic of the Randomization Group by Assessment interaction (Control and Baseline coded as the reference groups) as predictors of MCCB Domain scores and Composite.

b Effect sizes calculated as (Mean Change_{treatment} - Mean Change_{control})/pooled standard deviation of change. Note that effect size calculations include only subjects who completed both assessment points. Small effect: Cohen's $d=0.2$; Medium effect: Cohen's $d=0.5$; Large effect: Cohen's $d=0.8$ (Cohen, 1988)

* $p<.05$  § $p \leq .10$
Assessed for eligibility (n=201)
- Excluded (n=117)
  - Not meeting inclusion criteria (n=87)
  - Declined to participate (n=13)
  - Other reasons (n=17)

Enrolled (n=84)
- Discontinued (n=9)
  - Symptom exacerbation (n=3); time commitment (n=4); moved (n=2)

Randomized (n=75)

Allocated to intervention (n=39)
- Received allocated intervention (n=39)
- Did not receive allocated intervention (n=0)

Allocated to control (n=36)
- Received allocated intervention (n=33)
- Did not receive allocated intervention (n=3; discontinued prior to initiation)

Lost to follow-up (n=4)
- Symptom exacerbation (n=2), unable to contact (n=2)
- Discontinued intervention (n=6)
- Time (n=5), dissatisfaction with games (n=1)

Lost to follow-up (n=5)
- Symptom exacerbation (n=3), unable to contact (n=2)
- Discontinued intervention (n=4)
- Time commitment (n=4)

Lost to follow-up (n=0)
- Discontinued intervention (n=7)
- Time (n=6), frustration with games (n=1)

Lost to follow-up (n=4)
- Symptom exacerbation (n=2), unable to contact (n=2)
- Discontinued intervention (n=3)
- Time (n=2), moved (n=1)

Lost to follow-up (n=1)
- Unable to contact (n=1)
- Discontinued intervention (n=0)

Lost to follow-up (n=2)
- Unable to contact (n=2)
- Discontinued intervention (n=0)

Analysed (n=39)
- Excluded (n=0)

Analysed (n=33)
- Enrolled but discontinued prior to initiating training (n=3)
Cognitive change in T scores from Baseline to Post-Treatment (T3) and Durability (T4) by randomization group. ProcSpeed = Processing Speed; Attn = Attention; Wmem = Working Memory; Verbal = Verbal Learning and Memory; Visual = Visual Learning and Memory; ProbSolv = Problem Solving; Social = Social Cognition; CR = Cognitive Remediation; T3 = Post-Treatment Assessment; T4 = Durability Assessment. Data are based on change scores, and therefore include only participants who completed both relevant assessments.