

Reliability and Validity of THINC-it in Evaluating Cognitive Function of Patients with Bipolar Depression

This article was published in the following Dove Press journal:
Neuropsychiatric Disease and Treatment

Weihua Zhang,^{1,2,*} Na Zhu,^{3,*}
Jianbo Lai,^{1,4} Jingjing Liu,^{1,5}
Chee H Ng,⁶ Jun Chen,³
Chao Qian,⁷ Yanli Du,¹
Chanchan Hu,^{1,4} Jingkai Chen,^{1,4}
Jianbo Hu,^{1,4} Zhong Wang,^{1,4}
Hetong Zhou,^{1,4} Yi Xu,⁸⁻¹¹
Yiru Fang,³ Chuan Shi,⁸⁻¹¹
Shaohua Hu^{1,4}

¹Department of Psychiatry, First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, People's Republic of China; ²Department of Psychiatry, Taizhou Second People's Hospital, Taizhou 317200, People's Republic of China; ³Shanghai Mental Health Center, Shanghai Jiaotong University School of Medicine, Shanghai 200030, People's Republic of China; ⁴The Key Laboratory of Mental Disorder Management of Zhejiang Province, Hangzhou 310003, People's Republic of China; ⁵Department of Psychiatry, Wenzhou Kangning Hospital, Wenzhou 325000, People's Republic of China; ⁶The Melbourne Clinic Department of Psychiatry, University of Melbourne, Melbourne, Victoria 3052, Australia; ⁷Department of Psychiatry, The Seventh Shaoxing People's Hospital, Shaoxing 312000, People's Republic of China; ⁸Department of Psychological Assessment, Peking University Sixth Hospital, Beijing 100191, People's Republic of China; ⁹Peking University Institute of Mental Health, Beijing 100191, People's Republic of China; ¹⁰NHC Key Laboratory of Mental Health, Beijing 100191, People's Republic of China; ¹¹National Clinical Research Center for Mental Disorders, Beijing 100191, People's Republic of China

*These authors contributed equally to this work

Correspondence: Yiru Fang
Shanghai Mental Health Center, Shanghai Jiaotong University School of Medicine, Shanghai 200030, People's Republic of China
Email yirufang@aliyun.com

Shaohua Hu
Department of Psychiatry, First Affiliated Hospital, Zhejiang University School of Medicine, The Key Laboratory of Mental Disorder Management of Zhejiang Province, 79 Qing Chun Road, Hangzhou 310003, People's Republic of China
Email dorhushaohua@zju.edu.cn

Purpose: The THINC-integrated tool (THINC-it) as a brief screening tool can assesses cognitive impairment in patients with major depressive depression (MDD). Here, we aim to evaluate the reliability and validity of the THINC-it in a bipolar depression (BD-D) group in comparison with a healthy control (HC) group.

Materials and Methods: Both groups were matched according to age, gender, years of education, and IQ. All participants completed the THINC-it test, including Spotter, Symbol Check, Codebreaker, Trails, and the Perceived Deficits Questionnaire for Depression-5-item (PDQ-5-D). The concurrent validity and internal consistency of the THINC-it test were analyzed, and 30 healthy controls were randomly sampled to retest THINC-it to verify the reliability of the THINC-it retest. The correlation between THINC-it and Hamilton Depression Scale (HAMD-17) and Sheehan Disability Scale (SDS) was also analyzed.

Results: Fifty-eight patients with BD-D and 61 HCs were included for final analysis. There were significant mean difference (MD) standard errors (SE) between two groups in PDQ-5-D, Spotter and Codebreaker (all $P < 0.01$), Trails ($P = 0.015$). There was no significant difference in Symbol Check (MD (SE) = -0.01 (0.18), $P = 0.938$; 95% CI = -0.38 to 0.35). The Cronbach's α of PDQ-5-D was 0.640. The intraclass correlation coefficient (ICC) was between 0.440 and 0.757. The highest concurrent validity was PDQ-5-D ($r = 0.812$, $P < 0.001$). PDQ-5-D was positively correlated with HAMD-17 and SDS score ($P < 0.01$). The objective test had no significant correlation with HAMD-17 and SDS scores ($P > 0.05$).

Conclusion: This study found that THINC-it can accurately present the cognitive impairment of patients with BD-D. It can be potentially applied in assessing the cognitive function of patients with BD-D although Symbol Check may not accurately reflect the level of cognitive function. The concurrent validity and retest reliability are lower than expected, we need to further increase the sample size to study.

Keywords: bipolar depression, BD-D, major depressive depression, MDD, THINC-it, cognitive function, reliability, validity

Introduction

Bipolar disorder (BD) is a chronic mental disorder characterized by manic, hypomanic and depressive episodes.¹ Over the past decade, it has been demonstrated that bipolar depression (BD-D) patients have significant cognitive deficits in several cognitive domains, including planning, operational memory, attention, problem-solving, inhibition control, and psychological flexibility, which result in poor social functioning.² Studies have shown that cognitive function (especially executive

function and motor ability) is seriously impaired in patients with bipolar depression.³ Further, compared with major depressive depression (MDD), bipolar depression has a lower onset age, more mood episodes, more hospitalizations, and higher incidence of psychotic features. Numerous studies have shown that patients with bipolar depression at any age can have cognitive impairment, even in remission.⁴⁻⁹

The cognitive impairment of bipolar depression patients is significantly greater than that of patients with MDD.¹⁰ The causes of cognitive impairment in BD-D patients are not clearly elucidated. Multiple factors related to cognitive impairment have included changes of hippocampal structure and function (neural circuit),¹¹ obesity,¹² insufficient or excessive sleep,^{13,14} childhood adversity,¹⁵ medications,¹¹ immune dysfunction, and increased inflammatory cytokines.¹⁶ Patients with a higher level of education may have reduced learning and memory deficits,⁹ while a high level of cultural knowledge plays a protective role in the cognitive function of patients with bipolar depression.¹⁷

To address the cognitive outcomes of BD patients and reducing disease-related disability, effective assessment of cognitive impairment in BD patients is needed.⁸ Although HAMD and MADRS are used to assess depressive symptoms, both scales show minimal correlation with objective cognitive measurement,¹¹ and thus would rely on subjective clinical assessment and patient self-reports to evaluate cognitive function.⁹ Other scales used to evaluate cognitive function include: MCCB, which focuses on clinical tests of cognitive function in schizophrenic patients,^{18,19} and MoCA, which is often used as a general cognitive screening and dementia assessment tool.²⁰ MMSE and MoCA are suitable for elderly patients but generally not for young patients.²¹ In 2010, the International Society for Bipolar Disorders (ISBD) suggested that five subtests of MCCB, the California Verbal Learning Test, Stroop Test, Trail Making Test–part B, and Wisconsin Card Sorting Test can be used to test cognitive function of patients with BD.²² The Brief Assessment of Cognition in Affective Disorders (BAC-A) can also be applied to assess the global function of patients with bipolar disorder.²³ However, these scales are paper-based tests, which are time-consuming and laborious. Due to the lack of suitable brief and quick measurement tools, measurements of cognitive function remain a challenge.²⁴

Currently, there is no “gold standard” tool to evaluate or measure the cognitive function for the patients with

bipolar disorder.⁸ Therefore, a simple, easy to use and accurate cognitive impairment assessment which includes both subjective and objective evaluation, and good validity is needed.

The THINC-integrated tool (THINC-it) was released at the 2016 annual meeting of the European Society of Neuropsychopharmacology (ECNP) (website: <https://thinc.progress.im/en>). It evaluates the cognitive function of MDD and can be applied electronically on portable tablets. The test consists of Spotter, Symbol Check, Codebreaker, Trails, and PDQ-5-D. This set of assessment tools has been validated internationally and been proven to be an effective and sensitive tool for detecting adult MDD with cognitive dysfunction.²⁵ The time stability, reliability, and convergence effectiveness of THINC-it have also been shown in normal subjects.²⁶ In a study of 100 patients with generalized anxiety disorder (GAD) and 100 healthy controls, THINC-it effectively evaluated cognitive dysfunction of GAD patients where linear regression analysis showed that anxiety symptoms accounted for 70.4% of the variability of subjective cognitive disorder.²⁷ A recent study showed that the cognitive defects identified by THINC-it were correlated with overall and specific psychosocial defects, thus confirming the clinical value and utility of THINC-it as a cognitive screening tool for MDD patients.²⁸

The THINC-it is a self-administered tool which requires some instructions given to the subject who must be able to operate a computer independently. Objective and subjective methods are used to screen cognitive impairment of patients with depression,⁸ including cognitive tests of attention, working memory, processing speed, executive function, and cognitive function of capturing subjective perception.²⁹

THINC-it is recommended for early detection of cognitive impairment in mood disorders including BD-D.³⁰ However, there is currently no evaluation of THINC-it in bipolar depression. Therefore, we evaluated the reliability and validity of THINC-it to measure the neurocognitive impairment in a bipolar depression (BD-D) group compared with a healthy control (HC) group.

Materials and Methods

Subjects

All procedures in this study adhered to the Declaration of Helsinki (2008) Ethical Principles for Medical Research Involving Human Subjects and were approved by the

ethics committee of Zhejiang University, and all study participants gave voluntary and written informed consent.

Patients with BD-D Group

Patients diagnosed with bipolar disorder, current episode depressive (according to DSM-5 criteria and confirmed on MINI questionnaire), ie, bipolar depression (BD-D), without any other physical and psychiatric comorbidities, were recruited from May 2018 to April 2019 in the department of psychiatry, the First Affiliated Hospital, Zhejiang University School of Medicine. The inclusion criteria were as follows: 1) Age range from 18–65 years, both male and female; 2) outpatients or newly admitted inpatients who had not received any new treatment; 3) The score of Hamilton Depression Scale (HAMD-17) was no less than 17; 4) Mood Disorder Questionnaire (MDQ) score was positive (more than 7 items); 5) $IQ \geq 80$ on the Simple Wechsler Adult Intelligence; and 6) Subjects who had signed informed consent and were able to understand the self-report questionnaire. The exclusion criteria were as follows: 1) Alcohol and/or substance use disorders, autism spectrum disorder, dementia or any other neurodegenerative disorders, learning disabilities, schizophrenia or other mental disorders, or other medical conditions that may affect cognitive function, eg, brain tumors, multiple sclerosis, Parkinson's disease, etc.; 2) Patients on any drugs which may affect cognitive function, such as glucocorticoids, beta-blockers, opioid analgesics, central stimulants, etc.; 3) Patients who took benzodiazepines within 12 hours or alcohol within 8 hours of THINC-it administration; or 4) Patients who received electroconvulsive therapy (ECT) within the last 6 months.

Each participant completed the following: five tests of THINC-it, five criteria-related tests, the Wechsler Intelligence Test, HAMD-17, and the Sheehan Disability Scale.

Healthy Controls (HC) Group

Healthy controls were voluntarily recruited from Shangcheng District of Hangzhou City, Zhejiang Province, including junior high school and senior high school students, nurses, security guards, and social workers. They were aged 18–65, male or female. The MINI questionnaire confirmed that they did not meet any diagnostic criteria of previous psychiatric disorders. They also did not have any history of neurological diseases, physical diseases, alcohol dependence or smoking habits, or family history of psychiatric disorder among their first-degree

relatives. In total, 80 healthy participants were recruited in the study. The exclusion criteria were the same as that of the BD-D group. Thirty healthy participants were randomly selected 1 week later to retest five items of THINC-it.

Cognitive function evaluation was assessed by two psychiatrists. Before the evaluation, the consistency of scale scores was tested and no significant difference between raters were found. As far as possible training related procedures were adopted to eliminate any errors caused by subjective factors.

Measurements

The general assessment tools include the iPad version of THINC-it tool. The THINC-it complete set of tests includes Spotter, Symbol Check, Codebreaker, Trails, and the Perceived Deficits Questionnaire for Depression-5-item (PDQ-5-D).

The corresponding five criterion tests are the Perceived Deficits Questionnaire for Depression (PDQ-D), Reaction Time Paradigm (RTI), Digit Symbol Substitution Test (DSST), Trail Making Test-Part-B (TMT-B), and One-Back Task (1-back) (see Table 1). PDQ-D, DSST, and TMT-B are Pen-and-Paper Tests, while RTI is an online computer test, which is a subtest of the Cambridge Neuropsychological Test Automated Battery (CANTAB), the reaction time is taken as the result of this test, 1-back is also a computerized test developed by E-prime programme. The Sheehan Disability Scale (SDS) was then assessed. The Chinese version of PDQ-D is effective in psychometric evaluation of subjective cognitive impairment in MDD patients with good internal consistency and reliability of retest.³¹ The Sheehan Disability Scale is a self-rated disability scale, which evaluates three areas of disability: work, social life, and family life, with sufficient reliability and validity.³²

Each participant took about 15 minutes to complete five tests. Considering that different test sequences may have a potential impact on the results of the THINC-it test and Pen-and-Paper test, all participants were tested in the following order. Firstly, the demographic registration was completed, then the HAMD-17 tests were assessed. The MDQ was used to determine whether the participants met the diagnosis of BD-D. Four of the Wechsler intelligence tests were carried out, including common sense questions, arithmetic, similarity tests, number sequencing tests, and the total score was converted to IQ values. Then the objective test items PDQ-D, DSST, TMT-B, 1-back, and

Table 1 Description of Cognitive Tasks in THINC-it Tests and the Criterion Tests

Subtest	Definition	Direction	Outcome
PDQ-5-D	The questionnaire includes the five questions from the Perceived Deficits Questionnaire that evaluate attention/concentration and planning/ organization, as well as retrospective and prospective memory. The patient should select the appropriate response for each question considering their experiences during the previous 7 days.	↓	Sum of items
Spotter	Assessment of attention and executive function takes 2 minutes to complete. A series of arrows will appear, facing left or right. The user needs to choose the left or right direction as soon as possible according to the direction of the arrow. Their reaction causes the next arrow to appear. a reaction time test of attention.	↓	Mean of the log-transformed reaction time (seconds)
Symbol Check	The Symbol Check evaluating working memory, executive function, and attention/ concentration. Users are presented with a laterally moving sequence of symbols, the first of which is then hidden. The patient must correctly recall the hidden symbol as quickly as possible. Users are scored on speed and accuracy.	↑	Accuracy of trails
Codebreaker	Ability to identify executive functions, processing speed, and attention/ concentration, Six symbols are numbered consecutively. Users are presented with a series of numbers that they must then match with the correct corresponding symbol as quickly as possible.	↑	Total number correct
Trails	Chinese characters (壹贰叁肆伍陆柒捌玖) and Numbers (1–9) are presented on the screen. Subjects are connected in the sequence of 壹-1-贰-2-叁-3-肆-4-伍-5-陆-6-柒-7-捌-8-玖-9. The shorter the time used, the better the cognitive function, the longer the completion time, and the worse the cognitive function.	↓	Time to complete (seconds)
PDQ-D	PDQ-D is a questionnaire for MDD patients which assesses attention and concentration, prospective memory, retrospective memory, planning, and organization. There are 20 items in the questionnaire, with scores ranging from 0–4. The total score of PDQ-D is composed of the original scores of these 20 items, so it can range from 0–80. The score is positively associated with the severity of the cognitive impairment.	↓	Sum of items
RTI	The standard deviation calculated between the onset of the stimulus and the time at which the subject released the button. Calculated for correct, assessed trials in which the stimulus could appear in one location only.	↓	Mean of the log-transformed reaction time (seconds)
I-back	The one-back test asks the participants to determine whether the position of the second stimulus is the same as that of the previous stimulus, and then the subject chooses the right or left mouse button. It mainly tests the attention, memory, and reaction speed of the subjects, including the executive function.	↑	Accuracy of trails
DSST	DSST is based on a number-symbol encoding table, where each number of 1–9 corresponds to a symbol. Participants matched symbols and numbers as soon as possible in 90 seconds according to the encoding table. The higher the score is, the better the cognitive function is.	↑	Total number correct
TMT-B	TMT-B tests executive function, and subjects connect in the order of crossing numbers and letters. The higher the score, the more significant the functional impairment. TMT-B assesses the executive function of the subjects.	↓	Time to complete (seconds)

Note: Spotter: ..., not applicable.

Abbreviations: THINC-it, THINC-integrated tool; PDQ-5-D, Perceived Deficits Questionnaire for Depression-5-item; PDQ-D, Perceived Deficits Questionnaire for Depression; TMT-B, Trail Making Test-Part B; I-back, one-back test; DSST, digit symbol substitution test; RTI, reaction time paradigm.

RTI were completed. Five items of THINC-it were then tested. The THINC-it test was conducted in the order of PDQ-5-D, Spotter, Symbol Check, Codebreaker, and Trails. The BD-D group and the healthy control group completed the whole items in this order, and tried to achieve the same test link to reduce errors. Before each test, detailed instructions were given to the subjects. All five tests of THINC-it needed to be completed continuously in a single session. The order and content of the BD-D group and healthy control group were the same, so as to minimize errors.

Statistical Analysis

All data were analyzed by IBM SPSS Statistics 23.0 software package, including group data *t*-test for independent samples, chi-square test for classified data, reliability test, internal consistency test, concurrent validity test, and correlation analysis with clinical symptoms. The data that support the findings of this study are available on request from the corresponding author.

In the various tests of THINC-it, the score of Spotter and Trails were positively associated with the severity of the cognitive impairment. In terms of Symbol Check and Codebreaker, on the contrary, the original score is negatively associated with the severity of the cognitive function. In the original THINC-it design, the direction of each test result should be consistent. In statistical data, PDQ-5-D, Spotter and Trails are converted into standard *Z* score and multiplied by -1 , so the results of five tests indicate that the higher the score, the better the cognitive function. The data of PDQ-D, RTI, and TMT-B were analyzed by the same methods.

Results

A total of 98 patients with BD-D and 80 healthy controls were recruited. In the BD-D group, 21 patients were under the age of 18, 13 patients did not complete the test, and one patient had bipolar disorder, current episode mixed, five patients had $IQ < 80$. In the healthy control group, two participants were under 18 years old, one has a family history, and 16 participants did not complete the test. Finally, in total 58 patients with BD-D and 61 healthy controls were included in the analyses.

The BD-D and healthy control groups completed all cognitive function assessment in the first test, signed informed consent, including general demographic data, THINC-it complete set, five criterion tests. A week later, they were randomly retested, with 30 healthy controls

included in the retest. In terms of demographic data, there were no statistically significant group differences in age, gender, years of education, and IQ ($P > 0.05$). In the BD-D group the HAMD-17 score was 23.29 ± 5.20 and Sheehan Disability Scale score was 14.50 ± 6.75 . The age of onset in BD-D group was 20.33 ± 7.94 , and the duration of the disease was 203.56 ± 332.64 days. There was no statistical difference in IQ between two groups (see Table 2).

The *Z*-score was adopted to calculate the mean difference (SE) between two groups with THINC-it test. There were significant differences in PDQ-5-D, Codebreaker, Spotter, and Trails Mean difference (SE). The results are presented in Table 3. There was no significant difference in Symbol Check (MD (SE) = -0.01 (0.18), $P = 0.938$; 95% CI = -0.38 to 0.35). There were significant differences in *Z*-scores of THINC-it objective test (MD (SE) = 0.44 (0.12), $P = 0.001$; 95% CI = 0.19 to 0.69). There was also a significant difference in the *Z*-score of the total test (MD (SE) = 0.03 (0.10), $P = 0.748$; 95% CI = -0.17 to -0.24); The results of THINC-it test are basically consistent with those results of five criterion tests (see Table 3).

Table 2 Demographic and Clinic Characteristics of Subjects with BD-D Group and HC Group

Variables	BD-D (n=58)		HC (n=61)		P-value
	Mean	SD	Mean	SD	
Age, years	26.40	8.59	28.80	9.18	0.146
Education, years	13.83	2.67	14.56	2.81	0.149
IQ	124.47	18.77	128.66	20.55	0.249
HAMD-17 total score	23.29	5.20	0.48	1.36	<0.001**
SDS total score	14.50	6.75			
Age at onset of BD-D, years	20.33	7.94			
The course of the disease, days	203.56	±332.64			
Sex:	n=58	%	n=61	%	0.602
Male	22	37.93	26	42.62	
Female	36	62.07	35	57.38	

Notes: Spotter: . . . , not applicable; **P-value is significant at the 0.01 level.

Abbreviations: BD-D, bipolar depression; HC, healthy controls; IQ, intelligence quotient; HAMD, Hamilton Depression Scale; SDS, Sheehan Disability Scale; SD, standard deviation.

Table 3 Mean Difference in Performance on Individual Criterion Tests and Composite Z-Scores Between Subjects with BD-D Group and HC Group

	Measure	Mean Difference (SE)	P-value	95% CI
THINC-it	Spotter	-0.83 (0.17)	<0.001**	-1.16 to -0.49
	Symbol Check	-0.01 (0.18)	0.938	-0.38 to 0.35
	Codebreaker	0.51 (0.18)	0.005**	0.15 to 0.86
	Trails	-0.44 (0.18)	0.015*	-0.80 to -0.09
	PDQ-5-D	-1.59 (0.11)	<0.001**	-1.81 to -1.37
	Objective composite z-score	0.44 (0.12)	0.001**	-0.19 to 0.69
	Total composite z-score	-0.03 (0.10)	<0.001**	-0.17 to 0.24

Notes: Spotter: ..., not applicable; *P-value is significant at the 0.05 level; **P-value is significant at the 0.01 level.

Abbreviations: THINC-it, THINC-integrated tool; BD-D, bipolar depression; HC, healthy controls; SD, standard deviation; PDQ-5-D, Perceived Deficits Questionnaire for Depression-5-item; 95% CI, 95% confidence interval.

Reliability and Validity Test Results of THINC-it

Concurrent Validity

We used five criterion tests as the criterion, THINC-it five tests as the corresponding validity analysis, and applied the Pearson correlation analysis. In the BD-D group, PDQ-5-D and PDQ-D had the highest concurrent validity ($r=0.812$, $P<0.001$), the Codebreaker and DSST had high concurrent validity too ($r=0.747$, $P<0.001$). It is suggested that the parallel validity of PDQ-5-D and Codebreaker is similar to that of the criterion test tool. The concurrent validity of Symbol Check and 1-back ($r=0.456$, $P=0.001$) were poor. Only 26 patients in the BD-D group completed the RTI test with poor concurrent validity of Spotter and RTI ($r=0.389$, $P=0.017$), Trails and TMT-B ($r=0.360$, $P=0.006$) were lowest. Concurrent validity of objective tests ($r=0.695$, $P<0.001$); Concurrent validity of the overall test ($r=0.662$, $P<0.001$) (see Table 4).

Internal Consistency

With regards to internal consistency of THINC-it, the Cronbach's alpha of PDQ-5-D was 0.640. The internal consistency of THINC-it five total tests (PDQ-5-D, Spotter, Symbol Check, Codebreaker, Trails) was 0.412; the Cronbach coefficient of four objective tests (Spotter, Symbol Check, Codebreaker, Trails) was 0.641; in the criterion tests, the Cronbach's alpha of PDQ-D was 0.939 because of many items. The internal consistency of PDQ-D was very high. The Cronbach's coefficient of the total test (PDQ-D, DSST, TMT-B, 1-back) was 0.378, and the Cronbach's coefficient of the criterion tests objective test (DSST, TMT-B, 1-back) was 0.521 (see Table 4). The results of internal consistency were poor.

Reliability Test of Retest

Thirty subjects in the healthy control group were randomly retested, and all subjects were retested with the THINC-it test 1 week after the first test. All retest subjects completed the task well, using intraclass correlation efficient (ICC) analysis. ICC was between 0.440 and 0.757, the Codebreaker ($ICC=0.757$, $P<0.001$) and Symbol Check ($ICC=0.737$, $P<0.001$) were acceptable repeatability. The PDQ-5-D ($ICC=0.683$, $P<0.001$), Spotter ($ICC=0.643$, $P<0.001$) and Trails ($ICC=0.440$, $P=0.008$) had poor repeatability, the $ICC=0.793$ ($P<0.001$) in THINC-it objective composite score, and $ICC=0.781$ in THINC-it total composite score (see Table 4).

The Correlation Between the Test Results with HAMD-17 and SDS

The correlation between the scores of the test and HAMD-17 was compared. The peak, skewness, and Shapiro-Wilk test of HAMD-17 did not follow normal distribution. As most of the items in THINC-it did not follow normal distribution, we used the Spearman coefficient to calculate the correlation. PDQ-5-D was positively correlated with HAMD-17 ($P<0.01$). There was no significant correlation between objective test of THINC-it and HAMD-17 score ($P>0.05$). In the correlation comparison with the total score of Sheehan Disability Scale (SDS), as the total score of SDS was not distributed normally through kurtosis, skewness calculation, and Shapiro-Wilk test, we used the Spearman coefficient analysis. PDQ-5-D was positively correlated with SDS ($P<0.01$). Similar to the previous HAMD-17 analysis results, there was no significant correlation in the objective tests of THINC-it and SDS score ($P>0.05$) (see Table 5).

Table 4 Psychometric Properties of THINC-it Among Subjects with BD-D Group

Concurrent Validity			
THINC-it Test	Criterion Tests	Pearson r	P-value
Spotter (n=58)	RTI (n=26)	0.389	0.017**
Symbol Check (n=58)	1-back (n=52)	0.456	0.001**
Codebreaker (n=58)	DSST (n=57)	0.747	<0.001**
Trails (n=58)	TMT-B (n=57)	0.360	0.006**
PDQ-5-D (n=58)	PDQ-D (n=58)	0.812	<0.001**
THINC-it objective composite score (n=51)	Criterion Tests objective composite score (n=58)	0.695	<0.001**
THINC-it total composite score (n=51)	Criterion Tests total composite score (n=58)	0.662	<0.001**
Internal Consistency			
Test	n	No. of Items	Cronbach's α
THINC-it PDQ-5-D	58	5	0.640
THINC-it total composite score	58	5	0.412
THINC-it objective composite score	58	4	0.641
Criterion Tests objective composite score	51	4	0.521
Criterion Tests total composite score	51	4	0.378
PDQ-D	58	20	0.939
Intraclass Correlation Coefficient (ICC)			
Variable (n=30)	ICC	95% CI	P-value
PDQ-5-D	0.683	0.430–0.836	<0.001**
Spotter	0.643	0.370–0.813	<0.001**
Symbol Check	0.737	0.515–0.866	<0.001**
Codebreaker	0.757	0.548–0.877	<0.001**
Trails	0.440	0.093–0.689	0.008**
THINC-it objective composite score	0.793	0.608–0.896	<0.001**
THINC-it total composite score	0.781	0.588–0.890	<0.001**

Notes: Spotter: ..., not applicable; **P-value is significant at the 0.01 level.

Abbreviations: THINC-it, THINC-integrated tool; PDQ-5-D, Perceived Deficits Questionnaire for Depression-5-item; PDQ-D, Perceived Deficits Questionnaire for Depression; 95% CI, 95% confidence interval; RTI, reaction time paradigm; TMT-B, Trail Making Test-Part B; 1-back, one-back test; ICC, intraclass correlation coefficient; DSST, digit symbol substitution test.

Discussion

In this study, we validated the THINC-it tool to measure cognitive function in a bipolar depression group and compared with a healthy control group. Except for Symbol Check, there were significant differences in cognitive function between the two groups in Spotter, PDQ-5-D, Codebreaker, and Trails, which thus indicates that THINC-it can effectively evaluate the cognitive dysfunction of BD-D.

The study showed that the internal consistency reliability (Cronbach's α) of the subjective test item PDQ-5-D of THINC-it was 0.640. The Cronbach coefficient of four objective tests was 0.641, and the total Cronbach coefficient of five objective tests was 0.412. Previous study of THINC-it had demonstrated the internal consistency reliability of PDQ-5-D (Cronbach's $\alpha=0.769$), four objective composite tests

(Cronbach's $\alpha=0.551$), and five total composite tests (Cronbach's $\alpha=0.370$)²⁷ are basically consistent with our test results.

PDQ-5-D and PDQ-D, Codebreaker, and DSST had high concurrent validity, but the other three subtests of THINC-it (including Symbol Check and 1-back, Spotter and RTI, Trails and TMT-B) did not show high concurrent validity, and the concurrent validity results were basically the same as that of THINC-it's previous internal tests.²⁷ In general, the concurrent validity was low. The reason may be that the criterion is similar to THINC-it, when the THINC-it test is completed, then the criterion test is done, equivalent to doing the same test twice. As the participants become more familiar with the test tools, the test scores are increased, and the results are different, thus affecting the concurrent validity value.

Table 5 The Correlation Between the Results of Various Tests and the Total Score of Clinical Symptoms

The correlation between the test results and the total score of HAMD-17 (n=58)			
	Measure	Spearman r	P-value
THINC-it Test	Spotter	0.015	0.389
	Symbol Check	0.090	0.501
	Codebreaker	-0.040	0.767
	Trails	0.047	0.727
	PDQ-5-D	0.416	0.001**
The correlation between the test results and the total score of SDS (n=58)			
	Measure	Spearman r	P-value
THINC-it Test	Spotter	-0.046	0.730
	Symbol Check	-0.095	0.477
	Codebreaker	-0.072	0.590
	Trails	-0.059	0.661
	PDQ-5-D	0.430	0.001**

Notes: Spotter: ..., not applicable; **P-value is significant at the 0.01 level.

Abbreviations: THINC-it, THINC-integrated tool; PDQ-5-D, Perceived Deficits Questionnaire for Depression-5-item; HAMD, Hamilton Depression Scale; SDS, Sheehan Disability Scale.

For the retest reliability test, we retested 30 subjects in the healthy control group randomly, Codebreaker (ICC=0.757) and Symbol Check (ICC=0.737) were acceptable repeatability. The value of PDQ-5-D (ICC=0.683) exceeded the minimum acceptable level of 0.65. Spotter (ICC=0.643) and Trails (ICC=0.440) had poor repeatability. In this test, ICC value is low, and the reason may be the same as repeated test, the second time more skilled than the first. In comparison with scores of clinical symptoms, subjective test PDQ-5-D was significantly correlated with HAMD-17 and SDS scores, which could well reflect the degree of depression of patients, The four objective tests of the BD-D group and HC group had no significant correlation with HAMD-17 and SDS scores ($P>0.05$). One study found that differences in workplace performance among MDD patients are largely determined by the level of cognitive impairment, not by the severity of depression.³³ Increasing evidence indicated that impaired cognitive function might be an independent symptom in mood disorder.^{34–37}

In one study, 262 unmedicated MDD patients (161 females and 101 males) were tested with 24-Hamilton Depression Rating Scale (HDRS) and Beck Depression Inventory. Cognitive performance was not significantly correlated with the standard depression severity scale commonly

used in clinical trials.³⁸ Similarly, there was no correlation between the Chinese version of PDQ-D and DSST.³¹

In the comparison of mean difference between the bipolar depression group and the healthy control group, there were significant differences in PDQ-5-D, Spotter, Codebreaker, and Trails ($P<0.05$). However, in terms of Symbol Check, there was no significant difference between the depression group and the control group ($P=0.938$), and the test effect on cognitive function was slightly worse, with poor correlation with clinical symptoms. The test results in our study are basically the same as those of the THINC-it project team.²⁵ When the THINC-it psychological measurement was conducted, it was found that the convergence validity of Symbol Check task was low.²⁶

In our study, in order to reduce the test error, bipolar depression group, and healthy control group completed the testing in a strict order. However, variability in educational levels, depression severity, and cognitive function resulted in differences in the duration (1.5–4 hours), quality, and accuracy (about 50%) in completing the test. One study pointed out that the Symbol Check test was not a reliable alternative to the 1-back test.²⁶ In interpreting the results, we feel that it is not reasonable to conclude that the duration of test completion indicates the level of cognitive function as accuracy of response also needs to be taken into account. Further, the Symbol Check difficulty coefficient is large, which is more likely to result in the low score rather than the BD-D cognitive function decline.

Some subjects in the BD-D group of this study had high IQ values, which may offset the effects of cognitive impairment. In contrast, people with IQ higher than the average of the population by more than one standard deviation are less likely to be detected as “impaired”, even if their performance is significantly lower than their IQ prediction level.^{5,39}

In this study, five tests of THINC-it were programmed on an iPad, which was convenient and fast for self-administration. However, the performance of the tests is also dependent on adequate explanation and understanding of the instructions on how to complete the iPad version. It may take a long time for some participants to fully understand the rules and apply the skills to operate an electronic device. Hence, THINC-it tools which are programmed on tablets, or smartphones or personal computers may be a limitation, especially for elderly patients.⁴⁰

Limitations

First, despite the large sample, many patients were not completely familiar with the rules, and coupled with

unstable operation of the software system, thus resulting in significantly lower numbers included in the final analyses. Second, some patients had previous hospital experience, and may have taken similar tests, which may have affected the test results. Third, some subjects in the healthy control group had exaggerated PDQ-5-D and PDQ-D scores, but the subjects had no significant depressive mood. Fourth, many subjects in the BD-D group were older, seldom used computers, and had difficulty in operating tablet computers, which could have affected the test results. Last, patients with bipolar depression were not completely medication free, which may have affected the test performance.

Conclusion

In the study, we validated that THINC-it has a high sensitivity in the evaluation of cognitive function in patients with BD-D, however, its reliability and concurrent validity are low, and it has no significant correlation with clinical symptoms. In the practice, it is suggested to apply THINC-it to evaluate the cognitive function of the patients with BD-D. In the future, we can amplify the sample size to validate the clinical application.

Disclosure

Prof. Dr. Chee H Ng reports travel support and honorarium for delivery of talk from Lundbeck and Pfizer, and travel and honorarium for consulting service from Janssen, outside the submitted work. The authors report no other potential conflicts of interest in this work.

References

1. V V G, Sayana P, O F P, et al. Postmortem evidence of brain inflammatory markers in bipolar disorder: a systematic review. *Mol Psychiatry*. 2019;1–20. doi:10.1038/s41380-019-0448-7
2. M J M, Agid Y, Brüne M, et al. Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. *Nat Rev Drug Discov*. 2012;11(2):141–168. doi:10.1038/nrd3628
3. G S M, Ivanovski B, Hadzi-Pavlovic D, et al. Neuropsychological deficits and functional impairment in bipolar depression, hypomania and euthymia. *Bipolar Disord*. 2007;9(1–2):114–125. doi:10.1111/j.1399-5618.2007.00324.x
4. McIntyre RS, Soczynska JK, Mancini D, et al. Comparing features of bipolar disorder to major depressive disorder in a tertiary mood disorders clinic. *Ann Clin Psychiatry*. 2007;19(4):313–317. doi:10.1080/10401230701653617
5. Robinson LJ, Nicol Ferrier I. Evolution of cognitive impairment in bipolar disorder: a systematic review of cross-sectional evidence. *Bipolar Disord*. 2006;8(2):103–116. doi:10.1111/j.1399-5618.2006.00277.x
6. Samamé C, Martino DJ, Strejilevich SA. A quantitative review of neurocognition in euthymic late-life bipolar disorder. *Bipolar Disord*. 2013;15(6):633–644. doi:10.1111/bdi.12077
7. Jiménez E, Solé B, Arias B, et al. Impact of childhood trauma on cognitive profile in bipolar disorder. *Bipolar Disord*. 2017;19(5):363–374. doi:10.1111/bdi.12514
8. Raguett R-M, Cha DS, Kakar R, Rosenblat JD, Lee Y, McIntyre RS. Assessing and measuring cognitive function in major depressive disorder. *Evid Based Ment Health*. 2016;19(4):106–109. doi:10.1136/eb-2016-102456
9. Russo M, Mahon K, Burdick KE. Measuring cognitive function in MDD: emerging assessment tools. *Depress Anxiety*. 2015;32(4):262–269. doi:10.1002/da.22297
10. Bo Q, Dong F, Li X, et al. Comparison of cognitive performance in bipolar disorder, major depressive disorder, unaffected first-degree relatives, and healthy controls. *Psychiatry Clin Neurosci*. 2019;73(2):70–76. doi:10.1111/pcn.12797/full
11. McIntyre RS, Lee Y. Cognition in major depressive disorder: a “systemically important functional index” (SIFI). *Curr Opin Psychiatry*. 2016;29(1):48–55. doi:10.1097/YCO.0000000000000221
12. Cserjési R, Luminet O, Poncelet AS, Lénárd L. Altered executive function in obesity. Exploration of the role of affective states on cognitive abilities. *Appetite*. 2009;52(2):535–539. doi:10.1016/j.appet.2009.01.003
13. Ferrie JE, Shipley MJ, Akbaraly TN, Marmot MG, Kivimäki M, Singh-Manoux A. Change in sleep duration and cognitive function: findings from the Whitehall II study. *Sleep*. 2011;34(5):565–573. doi:10.1093/sleep/34.5.565
14. Kronholm E, Sallinen M, Suutama T, et al. Self-reported sleep duration and cognitive functioning in the general population. *J Sleep Res*. 2009;18(4):436–446. doi:10.1111/j.1365-2869.2009.00765.x
15. Lemos-Miller A, Kearney CA. Depression and ethnicity as intermediary variables among dissociation, trauma-related cognitions, and PTSD symptomatology in youths. *J Nerv Ment Dis*. 2006;194(8):584–590. doi:10.1097/01.nmd.0000230407.12687.ba
16. Rosenblat JD, Brietzke E, Mansur RB, Maruschak NA, Lee Y, McIntyre RS. Inflammation as a neurobiological substrate of cognitive impairment in bipolar disorder: evidence, pathophysiology and treatment implications. *J Affect Disord*. 2015;188:149–159. doi:10.1016/j.jad.2015.08.058
17. Grande I, Sanchez-Moreno J, Solé B, et al. High cognitive reserve in bipolar disorders as a moderator of neurocognitive impairment. *J Affect Disord*. 2017;208:621–627. doi:10.1016/j.jad.2016.10.012
18. Nuechterlein KH, Green MF, Kern RS, et al. The MATRICS consensus cognitive battery, part I: test selection, reliability, and validity. *Am J Psychiatry*. 2008;165(2):203–213. doi:10.1176/appi.ajp.2007.07010042
19. Green MF, Nuechterlein KH. The MATRICS initiative: developing a consensus cognitive battery for clinical trials. *Schizophr Res*. 2004;72(1):1–3. doi:10.1016/j.schres.2004.09.006
20. Davis DH, Creavin ST, Yip JL, Noel-Storr AH, Brayne C, Cullum S. Montreal cognitive assessment for the diagnosis of Alzheimer’s disease and other dementias. *Cochrane Database Syst Rev*. 2015;10:CD010775. doi:10.1002/14651858.cd010775.pub2
21. Baune BT, Miller R, McAfoose J, Johnson M, Quirk F, Mitchell D. The role of cognitive impairment in general functioning in major depression. *Psychiatry Res*. 2010;176(2–3):183–189. doi:10.1016/j.psychres.2008.12.001
22. Yatham LN, Torres IJ, Malhi GS, et al. The international society for bipolar disorders-battery for assessment of neurocognition (ISBD-BANC). *Bipolar Disord*. 2010;12(4):351–363. doi:10.1111/j.1399-5618.2010.00830.x
23. Bauer IE, Keefe RSE, Sanches M, et al. Evaluation of cognitive function in bipolar disorder using the brief assessment of cognition in affective disorders (BAC-A). *J Psychiatr Res*. 2015;60:81–86. doi:10.1016/j.jpsychires.2014.10.002
24. McIntyre RS, Lee Y, Carmona NE, et al. Characterizing, assessing, and treating cognitive dysfunction in major depressive disorder. *Harv Rev Psychiatry*. 2018;26(5):241–249. doi:10.1097/HRP.0000000000000171

25. McIntyre RS, Best MW, Bowie CR, et al. The THINC-integrated tool (THINC-it) screening assessment for cognitive dysfunction: validation in patients with major depressive disorder. *J Clin Psychiatry*. 2017;78(7):873–881. doi:10.4088/JCP.16m11329
26. Harrison JE, Barry H, Baune BT, et al. Stability, reliability, and validity of the THINC-it screening tool for cognitive impairment in depression: a psychometric exploration in healthy volunteers. *Int J Methods Psychiatr Res*. 2018;27(3):e1736. doi:10.1002/mpr.1736
27. Cha DS, Carmona NE, Rodrigues NB, et al. Cognitive impairment as measured by the THINC-integrated tool (THINC-it): the association with self-reported anxiety in major depressive disorder. *J Affect Disord*. 2018;238:228–232. doi:10.1016/j.jad.2018.05.006
28. Knight MJ, Fourrier C, Lyrtzis E, et al. Cognitive deficits in the THINC-integrated tool (THINC-it) are associated with psychosocial dysfunction in patients with major depressive disorder. *J Clin Psychiatry*. 2018;80(1):18m12472. doi:10.4088/JCP.18m12472
29. Baune BT, Malhi GS, Morris G, et al. Cognition in depression: can we THINC-it better? *J Affect Disord*. 2018;225:559–562. doi:10.1016/j.jad.2017.08.080
30. Atique-Ur-Rehman H, Neill JC. Cognitive dysfunction in major depression: from assessment to novel therapies. *Pharmacol Ther*. 2019;202:53–71. doi:10.1016/j.pharmthera.2019.05.013
31. Shi C, Wang G, Tian F, et al. Reliability and validity of Chinese version of perceived deficits questionnaire for depression in patients with MDD. *Psychiatry Res*. 2017;252:319–324. doi:10.1016/j.psychres.2017.03.021
32. Leon AC, Shear MK, Portera L, Klerman GL. Assessing impairment in patients with panic disorder: the sheehan disability scale. *Soc Psychiatry Psychiatr Epidemiol*. 1992;27(2):78–82. doi:10.1007/BF00788510
33. McIntyre RS, Soczynska JZ, Woldeyohannes HO, et al. The impact of cognitive impairment on perceived workforce performance: results from the International mood disorders collaborative project. *Compr Psychiatry*. 2015;56:279–282. doi:10.1016/j.comppsy.2014.08.051
34. Chaves OC, Lombardo LE, Bearden CE, et al. Association of clinical symptoms and neurocognitive performance in bipolar disorder: a longitudinal study. *Bipolar Disord*. 2011;13(1):118–123. doi:10.1111/j.1399-5618.2011.00888.x
35. Levy B, Medina AM, Weiss RD. Cognitive and psychosocial functioning in bipolar disorder with and without psychosis during early remission from an acute mood episode: a comparative longitudinal study. *Compr Psychiatry*. 2013;54(6):618–626. doi:10.1016/j.comppsy.2012.12.018
36. Martinez-Aran A, Vieta E, Colom F, et al. Neuropsychological performance in depressed and euthymic bipolar patients. *Neuropsychobiology*. 2002;46(Suppl. 1):16–21. doi:10.1159/000068016
37. Martino DJ, Strejilevich SA, Marengo E, Ibañez A, Scápola M, Igoa A. Toward the identification of neurocognitive subtypes in euthymic patients with bipolar disorder. *J Affect Disord*. 2014;167:118–124. doi:10.1016/j.jad.2014.05.059
38. Keilp JG, Madden SP, Gorlyn M, Burke AK, Oquendo MA, Mann JJ. The lack of meaningful association between depression severity measures and neurocognitive performance. *J Affect Disord*. 2018;241:164–172. doi:10.1016/j.jad.2018.08.034
39. McIntyre RS, Cha DS, Soczynska JK, et al. Cognitive deficits and functional outcomes in major depressive disorder: determinants, substrates, and treatment interventions. *Depress Anxiety*. 2013;30(6):515–527. doi:10.1002/da.22063
40. Fiorillo A, Carpinello B, De SG, et al. Assessment and management of cognitive and psychosocial dysfunctions in patients with major depressive disorder: a clinical reView. *Front Psychiatry*. 2018;9:493. doi:10.3389/fpsy.2018.00493

Neuropsychiatric Disease and Treatment

Dovepress

Publish your work in this journal

Neuropsychiatric Disease and Treatment is an international, peer-reviewed journal of clinical therapeutics and pharmacology focusing on concise rapid reporting of clinical or pre-clinical studies on a range of neuropsychiatric and neurological disorders. This journal is indexed on PubMed Central, the 'PsycINFO' database and CAS, and

is the official journal of The International Neuropsychiatric Association (INA). The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/neuropsychiatric-disease-and-treatment-journal>