BIPOLAR DISORDERS
AN INTERNATIONAL JOURNAL OF PSYCHIATRY AND NEUROSCIENCES

Bipolar Disorders 2012: 14: 118–125

Brief Report

Deficits in tasks of executive functioning that mimic real-life scenarios in bipolar disorder

Torralva T, Strejilevich S, Gleichgerrcht E, Roca M, Martino D, Cetkovich M, Manes F. Deficits in tasks of executive functioning that mimic real-life scenarios in bipolar disorder.
Bipolar Disord 2012: 14: 118–125. © 2012 The Authors.
Journal compilation © 2012 John Wiley & Sons A/S.

Background: A growing body of evidence suggests that patients with bipolar disorder (BD) have cognitive impairments even during euthymic periods. The main cognitive domains affected are verbal memory, attention, and executive function. Nevertheless, some studies suggest that at least a subgroup of euthymic patients demonstrates intact executive functioning in classic neuropsychological tests, which could be due to the lack of real-life, or *ecological* validity.

Objective: In this study, we highlight the usefulness of incorporating more ecological tests of executive function in assessment batteries in order to detect specific cognitive deficits in BD patients with otherwise normal performance in standard executive tests.

Methods: Nineteen euthymic BD patients and 15 healthy controls completed a standard neuropsychological battery assessment and two experimental tasks (the Multiple Errands Test–Hospital Version and the Hotel Task) to measure executive functioning in highly demanding cognitive settings that mimic real-life scenarios.

Results: No significant differences were found between the groups' demographic variables. We found, as predicted, that the group of euthymic BD patients who had control-comparable performance in classic executive tasks showed important deficits in more ecological tasks of executive functioning of the type that mimic real-life scenarios.

Conclusions: Together, these data suggest that the inclusion of ecological tests in the assessment of BD patients can contribute to providing a more realistic cognitive profile of this patient population, which will undoubtedly allow for a better design of therapeutic and rehabilitation strategies that can help patients to minimize impact in real-life settings.

Teresa Torralva[a,b,c], Sergio Strejilevich[a,c], Ezequiel Gleichgerrcht[a,c], María Roca[a,c], Diego Martino[c], Marcelo Cetkovich[a,c] and Facundo Manes[a,c]

[a]Institute of Cognitive Neurology (INECO), Buenos Aires, Argentina, [b]University Diego Portales, Santiago, Chile, [c]Department of Psychiatry, Institute of Neurosciences, Favaloro University, Buenos Aires, Argentina

doi: 10.1111/j.1399-5618.2012.00987.x

Key words: bipolar disorder – ecological tests – executive functions

Received 23 November 2010, revised and accepted for publication 1 October 2011

Corresponding author:
Teresa Torralva
Institute of Cognitive Neurology (INECO)
Pacheco de Melo 1854/60 (1126)
Buenos Aires
Argentina
Fax: +54 (11) 4812-0010
E-mail: ttorralva@ineco.org.ar

Understanding the extent of cognitive impairments in people with bipolar disorder (BD) has become a strong focus of research. A growing body of evidence suggests that patients with BD have cognitive impairments even during euthymic periods. Independent meta-analyses have concluded that the main cognitive domains affected in remitted patients are verbal memory, attention, and executive function (1, 2). Moreover, it has been shown that there is a negative association between neurocognitive functioning and different measures of disability in both cross-sectional (3–5) and longitudinal (6–8) studies.

An important study that compared the cognitive function in patients with BD to that of patients with schizophrenia found a bimodal pattern for executive functions in the BD cohort, with many in the group indistinguishable from healthy controls or from the group with schizophrenia (9). Another pilot study found that only the BD patients with

low psychosocial functioning exhibited impairments in executive functioning, while none of the BD patients with high psychosocial functioning exhibited any impairment in this area (10). A recent study also reported that around 40% of euthymic patients with BD exhibited no differences in executive functioning measures when compared to healthy controls (11). Taken together, these findings seem to suggest that at least a subgroup of euthymic patients might demonstrate intact executive functioning in the classic neuropsychological evaluation.

It is possible that some of these findings may reflect a known limitation to the standard executive tasks (i.e., their lack of real-life or *ecological* validity). Because executive functions constitute a multidimensional concept that includes different cognitive processes such as working memory, inhibitory control, cognitive flexibility, set shifting, planning, and organization, it is not uncommon that patients with frontal dysfunction may exhibit planning impairment in real-life situations, despite demonstrating adequate performance on traditional assessment measures. Gioia and Isquith (12) highlighted the fact that in classic neuropsychological tests, the examiner provides the structure, organization, guidance, planning, and monitoring necessary for optimal performance, therefore transforming himself or herself in the patient's own executive system. Thus, traditional testing environments may fail to induce executive deficits, making the assessment of this cognitive domain particularly challenging. Consequently, classic executive tests may not be sufficiently sensitive (13–15) to detect subtle or specific deficits. In fact, we have recently demonstrated that some conditions, such as adult attention-deficit hyperactivity disorder (16) or early behavioral variant frontotemporal dementia (17), can present within normal scores on standard neuropsychological assessment, yet exhibit performance deficits on tasks that pose real-life demands. It has also been suggested that some traditional executive tests, such as the Wisconsin Card Sorting Test (WCST) and the verbal fluency test, are greatly dependent on fluid intelligence (18). This would imply that there is a need to include tests of frontal function that measure functioning beyond the influence of IQ.

In clinical practice, being able to quantify the extent of cognitive/executive dysfunction in BD patients is particularly important, both to gain further understanding of these disorders and to help to design appropriate rehabilitation strategies with the ultimate aim of reducing the negative impact of these deficits on patients' daily lives. Therefore, the aim of our investigation was to highlight the usefulness of incorporating more ecological tests of executive functioning in the neuropsychological assessment of patients with BD in order to detect specific or subtle cognitive deficits with otherwise normal performance in standard tests of executive functioning. Accordingly, we incorporated a laboratory multitasking task (the Hotel Task) and a 'real-life' multitasking task [Multiple Errands Test–Hospital Version (MET-HV)], both of which have been demonstrated to have excellent ecological validity – that is, they actually mimic real-life scenarios and have proven to be sensitive in the detection of frontal dysfunction in various neurological and psychiatric conditions (16–19), even in the absence of deficits in standard cognitive tests (17, 20). We also administered an extensive standard cognitive battery of tests to fully characterize participants' cognitive profiles. Based on our previous studies, we hypothesized that the assessment of BD patients with multitasking ecological executive tasks would detect executive impairments differently than it would through the administration of classic executive tests.

Methods

Participants

Participants were outpatients at the Institute of Cognitive Neurology (Buenos Aires, Argentina) who underwent a detailed examination of their psychiatric and neuropsychological profile. Diagnoses of BD (n = 19) were made according to DSM-IV using Structured Clinical Interview for DSM-IV (SCID) (21) and patients had to be euthymic [defined by the Hamilton Depression Rating Scale (HDRS) (22) ≤ 8, and the Spanish version of the Young Mania Rating Scale (YMRS) (23) ≤ 6] for at least eight weeks with no change of medication type or dose over a period of four months. Exclusion criteria were: prior history of substance abuse, history of mental retardation, neurological disease, or any clinical condition that could affect cognitive performance. Healthy controls (n = 15) were randomly recruited from a larger pool of volunteers who had neither a history of abuse of recreational drugs nor a history of neurological or psychiatric disease in themselves or a first-degree family member.

Materials and procedures

The study was initially approved and supervised by the ethics committee at the Institute of Cognitive

Neurology and all participants signed an informed consent in accordance with the Declaration of Helsinki for research with human subjects. Patients first completed a series of psychiatric questionnaires in order to confirm their diagnoses and establish a profile of clinical symptoms. Depressive BD symptoms were rated using the HDRS and the Beck Depression Inventory (BDI–II) scale (24), and the Spanish version of the YMRS was used to rate manic symptoms. Social, functional, and occupational functioning was measured in BD patients using the General Assessment of Functioning (GAF) scale (25), which is a widely used DSM-based tool that assesses global functioning.

Participants in both groups completed a thorough classic neuropsychological battery (CNB) assessment of: (i) estimated premorbid IQ with the Word Accentuation Test–Buenos Aires (WAT-BA) (26); (ii) overall cognitive status using the Addenbrooke's Cognitive Examination (ACE) (27); (iii) attention with the Forward Digits Span Task of the Wechsler Adult Intelligence Scale–III (WAIS-III) (28) and the Trail Making Test–part A (TMT-A) (29); (iv) verbal memory through the logical memory subtest of the Wechsler Memory Scale–Revised (WMS-R) (30) and non-verbal memory with the Rey Complex Figure Test (31); (v) language with the adapted version of the Boston Naming Test (32) for denomination and the adapted version of the Token Test (33) for comprehension; and (vi) executive functioning using the Backward Digits Span Test (30), the Trail Making Test–part B (TMT-B) (29), the letters and numbers ordering test (30), the modified version of the WCST (34), and the Frontal Assessment Battery (FAB) (35).

MET-HV. This test, which is frequently administered in hospitals, requires participants to carry out a number of tasks simulating 'real-life' situations where minor inconveniences can take place (36). While still in the hospital, the patient is given a card with four sets of simple tasks, totaling 12 subtasks. The first set requires participants to achieve six specific goals, which include purchasing three items (a bottle of soda, a postcard, and a letter), collecting an envelope from reception, using the internal telephone, and posting something to an external address. The second set involves obtaining and writing down pieces of information (the area code of Chivilcoy, Argentina; the price of a dinner menu; the last transfer shuttle schedule to Buenos Aires). In the third set, the participant is required to call the evaluator 20 minutes after the test has begun and state the time over the telephone. The final task requires the participant to inform the proctor when every task has been completed. Nine rules are clearly stated in the instruction sheet and the participant's behavior while carrying out the tasks is monitored by two observers. At the end of the test, each participant has to indicate on a ten-point scale how well they think they have done. Errors in this test were categorized as: (i) *inefficiencies*: where a more effective strategy could have been applied; (ii) *rule breaks*: where a specific rule (social, or one of the nine explicitly defined within the test) was broken; (iii) *interpretation failure*: where the requirements of a task had been misunderstood; (iv) *task failures*: where any of the 12 tasks had not been fully completed; and (v) *total fails*: the sum total of all the previous categories.

The Hotel Task. We adapted the task proposed by Manly et al. (37) for the rehabilitation of executive symptoms, while preserving its most important features. The task comprises six types of activity that a person would plausibly need to undertake while working in a hotel. The materials needed to perform these activities were arranged on a desk and randomly distributed between participants and sessions. The instructions were as follows: '*In this task you are asked to imagine that you are working in a hotel. Your manager is keen for you to try each of these five everyday activities during the next 15 minutes so that you can get a feel for the tasks – and make an informed estimate of how long each task would take to complete. Your main goal is to attempt each of these five tasks over the next 15 minutes. There are five main tasks. Each of the tasks may take longer than 15 minutes to complete on its own, so there is no way that you will be able to complete all of them. The most important thing is to try to do a little of each task – spending as much time on each as possible within the total time available*'. The details for each of the tasks were then described, and a written summary of the task was placed on top of the relevant materials. The tasks were: (i) *compiling individual invoices:* participants were provided with a collection of invoices that needed to be arranged by guest name; (ii) s*orting the charity collection:* the materials included a box containing 200 coins, which needed to be grouped by country of origin (Argentina, France, Italy, USA, and Hungary); (iii) *looking up telephone numbers:* participants were provided with a list of 34 local companies and asked to find and note down their telephone numbers using the regional Yellow Pages telephone directory; (iv) *sorting conference labels:* participants were provided with a pile of 100 labels, each with the name of a guest attending a conference; the pile was shuffled and

participants were asked to sort the cards into alphabetical order; and (v) *proofreading the hotel leaflet:* participants were asked to check a nine-page draft of a proposed new leaflet for the hotel for typographic and grammatical errors. The scoring of the Hotel Task for this experiment was as follows: (i) the number of main tasks attempted (out of five); and (ii) the time deviation; the optimal time allocation was three minutes per task and deviations (in sec) from this timeframe were calculated and totaled [for detailed scoring see (37)].

Both the CNB and the experimental tests were administered by two experienced neuropsychologists (TT and MR).

Statistical analysis

Demographic and clinical data were compared between the groups using independent sample *t*-tests or Mann–Whitney *U*-tests when equal variances could not be assumed. When analyzing categorical variables (e.g., gender, recognition), χ^2 tests for 2×2 contingency tables were used. Any demographic and clinical variables that differed significantly between the groups were used as covariates in an analysis of variance between the groups on the neuropsychological variables. Spearman's rank correlation coefficients were conducted in order to further investigate the relationship between neuropsychological variables. The α-value for all statistical tests was set at 0.05, two-tailed.

Results

Demographic and clinical variables

No significant differences were found between the groups for age ($t_{32} = -0.29$, $p = 0.77$), years of education ($U = 119.5$, $p = 0.43$), gender ($\chi^2 = 3.34$, df = 1, $p = 0.49$), or premorbid intellectual functioning ($U = 115.0$, $p = 0.35$). As expected, the BD group scored significantly higher than controls on the BDI-II ($U = 9.50$, $p < 0.001$), which measures the severity of mood symptoms at present (Table 1). Forty-two percent of patients were receiving antipsychotic medication, 31% were being treated with benzodiazepines, 84% were on antidepressant medication, and 84% were taking mood stabilizers.

Table 1. Demographic and clinical information for the bipolar disorder and healthy control groups

	Bipolar disorder (n = 19)	Healthy controls (n = 15)
Demographics, mean (SD)		
Age, years	45.6 (12.9)	43.9 (19.8)
Gender, male/female	12/7	8/7
Education, years	15.4 (3.0)	14.6 (3.0)
Clinical profile, mean (SD)		
WAT-BA	38.9 (2.8)	37.5 (4.4)
BDI-II[a]	13.8 (11.5)	0.9 (2.0)
Age at onset, years	30.1 (14.4)	
Duration, years	12.8 (9.6)	
No. depressive episodes	4.1 (2.8)	
No. manic episodes	2.5 (1.8)	
No. hospitalizations	0.7 (0.9)	
HDRS score	7.8 (7.5)	
GAF score	75.7 (13.9)	
YMRS score	2.9 (3.2)	

SD = standard deviation; WAT-BA = Word Accentuation Test–Buenos Aires; BDI-II = Beck Depression Inventory; HDRS = Hamilton Depression Rating Scale; GAF = Global Assessment of Functioning; YMRS = Young Mania Rating Scale.
[a]A significant difference was found at $p < 0.001$ between the bipolar disorder and healthy control groups.

CNB versus Executive and Social Battery (ESCB) performance

In order to control for mood symptoms, total score on the BDI-II was used as a covariate in the analysis of variance on the variables of both the CNB and the experimental tasks of the ESCB. As shown in Table 2, no significant differences were found between the groups in any of the variables of CNB included in the comprehensive assessment of patients. In fact, as revealed by the η^2 values, the between-group effect (i.e., BD versus control) explained a maximum of 15% of the variance on neuropsychological performance, and the average effect size was 0.06 (SD = 0.04), revealing that performance differences between the groups are substantially reduced when the severity of mood symptoms is statistically controlled.

By contrast, when performance between the groups was compared on the ESCB, a significant difference was found in the variables of both experimental tasks (Table 3). In particular, the BD group showed significantly more inefficiencies ($F_{1,31} = 29.4$, $p < 0.001$) and rule breaks ($F_{1,31} = 17.9$, $p < 0.001$) than controls on the MET-HV task. On the Hotel Task, the controls attempted to complete significantly more tasks than BD patients ($F_{1,31} = 14.1$, $p = 0.001$). In these cases, these variables explained 42–58% of the variance in performance between groups, revealing the higher capacity of the ESCB ecological tasks to detect executive deficits.

Correlations

Within the BD group, a significant correlation was found between the GAF and the number of tasks

Table 2. Classic neuropsychological battery performance for the bipolar disorder and control groups

	Bipolar disorder (n = 19)	Healthy control (n = 15)	*F*	p-value	η^2
General cognitive status					
ACE	94.3 (1.1)	94.7 (1.3)	0.5	0.82	< 0.001
Attention					
Digit Forward Span	6.9 (0.4)	6.9 (0.4)	0.001	0.97	0.066
TMT-A (sec)	33.8 (3.2)	35.3 (3.7)	0.08	0.78	< 0.01
Memory					
Logical memory					
Delayed	18.8 (2.3)	25.2 (2.7)	2.63	0.12	0.10
Recognition	17.4 (0.6)	17.8 (0.7)	0.12	0.73	< 0.01
RAVLT					
Immediate	46.2 (2.5)	52.9 (2.9)	2.39	0.136	0.098
Delayed	9.9 (1.3)	10.0 (1.5)	0.003	0.95	< 0.001
Recognition	11.9 (0.9)	14.7 (0.9)	2.91	0.10	0.12
Rey Complex Figure					
Immediate	34.1 (0.9)	35.3 (1.0)	0.66	0.43	0.029
Delayed	17.3 (2.2)	22.9 (2.6)	2.19	0.15	0.091
Language					
Boston	19.6 (0.1)	19.9 (0.2)	0.65	0.43	0.029
Token Test	26.9 (0.7)	25.4 (0.8)	1.80	0.19	0.076
Phonological fluency	16.9 (1.3)	17.9 (1.5)	0.25	0.62	0.011
Semantic fluency	18.7 (1.5)	21.6 (1.7)	1.42	0.25	0.061
Executive functions					
Digit Backwards Span	4.4 (0.4)	5.2 (0.5)	1.55	0.23	0.066
TMT-B (sec)	77.1 (14.7)	105.3 (17.1)	1.27	0.27	0.054
WCST					
Categories	5.6 (0.2)	5.7 (0.2)	0.41	0.53	0.018
Perseverative errors	1.2 (0.6)	1.9 (0.7)	0.42	0.52	0.019
FAB total score	17.7 (0.1)	17.9 (0.1)	1.01	0.33	0.039

Values are shown as estimated marginal mean (standard error) following analysis of covariance, with severity of mood symptoms as a covariate. ACE = Addenbrooke's Cognitive Examination; TMT-A = Trail Making Test–part A; RAVLT = Rey Auditory Verbal Learning Test; TMT-B = Trail Making Test–part B; WCST = Wisconsin Card Sorting Test; FAB = Frontal Assessment Examination.

Table 3. *Ecological* neuropsychological battery performance for the bipolar disorder and control groups

	Bipolar disorder (n = 19)	Healthy controls (n = 15)	*F*	p-value	η^2
MET-HV					
Inefficiencies	4.22 (0.6)	0.06 (0.4)	29.40	< 0.001	0.58
Rule breaks	3.94 (0.6)	0.63 (0.4)	17.90	< 0.001	0.46
Interpretation failures	0.38 (0.3)	0.24 (0.2)	0.12	0.73	0.006
Task failures	0.64 (0.3)	0.42 (0.2)	0.48	0.49	0.022
Hotel Task					
Tasks attempted	3.77 (0.2)	4.90 (0.1)	14.10	0.001	0.42
Time deviation (sec)	334.10 (59.7)	319.90 (42.9)	0.03	0.86	0.001

Values are shown as estimated marginal mean (standard error) following analysis of covariance, with severity of mood symptoms as a covariate. MET-HV = Multiple Errands Test–Hospital Version.

attempted on the Hotel Task ($r = 0.55$, $p = 0.04$). The Hotel Task also correlated significantly with Backward Digit Span ($r = -0.53$, $p = 0.02$) and with performance on the TMT-B ($r = -0.43$, $p = 0.05$). Yet, the GAF did not correlate significantly with any of the classic tasks of executive function. No other significant correlations were found between clinical variables and neuropsychological performance either on standard or ecological tasks.

Discussion

As predicted, we have shown that a group of euthymic BD patients who had control-comparable performance in classic executive tasks showed important deficits in more *ecological* tasks of executive functioning of the type that mimic real-life scenarios. When comparing our BD patients and controls on these ecological tests, significant

differences were found in several variables of both the MET-HV and the Hotel Task. In particular, the BD group showed significantly more inefficiencies and rule breaks than controls on the MET-HV task. This may reflect BD patients acting more impulsively, with no apparent planning, and poor organization of the tasks. On the Hotel Task, healthy controls attempted to complete significantly more tasks than BD patients, indicating that this patient population may have a worse capacity for planning, flexibility, and organization. The fact that some patients, who clearly have impairments in everyday life settings but show few or no deficits in a wide range of standard tests of executive functions (38), reveals a major problem in neuropsychology. This can be explained by many classic tests of cognitive functioning, especially those of executive functions related to their lack of real-life, ecological validity.

There are several reasons why standard laboratory-based tests might fail to detect these subtle deficits, making assessment of this cognitive domain particularly challenging. Consequently, these tests might not be sufficiently sensitive (13, 39, 40) to identify deficits in real-life executive functioning. If the patient's environment makes little demand on certain skills, executive deficits might have no impact on real-life settings. By contrast, minor executive deficits can actually become especially impairing in highly demanding environments. The development of assessment tools that mimic real-life scenarios should focus on the detection of the actual cognitive demands involved in everyday real-life settings. Moreover, Roca et al. (18) have recently suggested that some traditional executive tests, such as the WCST and phonological fluency, both of which are extensively used worldwide, are greatly dependent on fluid (*g*) intelligence scores. This would suggest that for an optimal assessment of the frontal lobes, one should include tests that can gauge performance level beyond the *sole* influence of fluid intelligence and that are of special significance for identifying impaired everyday activities. In the context of these findings, it is important to note, however, that the present study does not intend to minimize the usefulness of classic tests of executive functioning. These tests not only provide a wide array of quantitative information, but also can be a rich source of qualitative data, especially with regard to the generation of strategies to complete the tasks, which is a key component of the executive domain. Instead, the goals of the present work were, first, to provide an alternative set of tests that complement standard tasks when the latter fail to detect deficits reported by relatives, caregivers, and patients; and, second, to encourage researchers to look for ways to maximize the sensitivity of standard tests. Future research should be devoted to generating new ways of quantifying strategy generation on the TMT-B which might increase its capacity to detect subtle deficits otherwise unnoticed with 'total time' and 'number of errors' on this task.

Both experimental tasks administered in this study were selected for being *unstructured* or *ill-structured* in nature, for they present participants with many possible courses of action to choose from in order to complete the task under certain given rules. It has been reported (36, 41) that patients with 'strategy application disorder' failed in this kind of multitasking test, making significantly more errors than controls, being less efficient, breaking more rules, and misinterpreting instructions, even when presenting normal scores on classic executive tests.

Clinically speaking, the detection of these deficits through the use of ecological tests is crucial because it can be very frustrating for relatives and caregivers to be told that cognitive performance is intact when they report deficits in real-life executive functioning. The same holds true for the patients themselves, who tend to have relatively accurate view of their own cognitive impairments, and often find it puzzling to hear that their cognitive assessment is considered completely normal. Moreover, the use of ecological tests would be useful in high-risk studies, taking into account that relatives of patients might have had more subtle executive dysfunction.

A growing area of research is the relationship between patients' neuropsychological outcomes and their real functional status. Although there is a general consensus that cognition is a strong predictor of social and adaptive functions, there is a lack of consensus on which exactly are the best measures that reflect functional status. The approach of coming up with better ecological validity tests could potentially offer a better predictor of functionality. In the present study, we found that only a measure derived from the experimental tasks (the number of tasks attempted on the Hotel task) correlated significantly with the GAF. Performance on the Hotel Task also correlated significantly with two classic executive tests (i.e., Digit Span Backward and TMT-B). However, the two classic executive tests did not correlate significantly with the GAF. These findings indicate some important associations. On the one hand, performance in classic tests seems to be related to performance in ecological tasks, although the moderate correlation coefficients reveal that these tasks are definitely not measuring exactly the same

processes. On the other hand, these findings also reveal that global functioning may be more closely related to performance on ecological tests than on classical executive tasks, which supports the idea that laboratory-based settings may not always be able to reflect real-life deficits. Several past studies (7, 8, 42) have nonetheless shown an association between the GAF and standard cognitive measures. Our findings, however, partly replicate a recent study conducted by O'Shea et al. (43) showing no correlation between cognitive functioning and social or occupational functioning in a group of euthymic BD patients. One possibility for the lack of correlation between GAF and performance on standard cognitive tests might be that patients included in this study had a high level of functioning. Future studies should attempt to compare BD patients with high- and low-functioning scores on ecological executive tasks.

These results, naturally, must be interpreted in the context of some limitations. First, and as with almost all previous studies, all patients included here were taking psychotropic medications, and we cannot discount the influence of drugs on cognitive functioning. Another limitation of our study was the relatively small sample size. Nevertheless, this study represents one of the first attempts to address the issue of real-life executive deficits in BD, and the significant differences and relationships obtained were statistically robust. Moreover, we did find significant effects, and the study groups were much larger than those employed in the other relevant reports on BD. Furthermore, the power of this study was maximized by the clear operational definitions and strong inclusion criteria, the control for affective symptoms, and accurately paired patients with healthy controls according to age and gender, ensuring proper characterization for each group.

The clear relationship that the results of this study establish between real-life functioning and performance on highly demanding tests such as the Hotel Task, creates a new neuropsychological framework in approaching BD. Tasks of this type might not only be useful in assessing executive functions in a more ecological way, but they may also inspire new cognitive rehabilitation programs that could be effective in training psychiatric patients to achieve better performance in their real lives.

To our knowledge, this was one of the first studies to examine a group of BD patients who showed control-comparable performance on standard neuropsychological assessment using novel and more ecological tasks of executive functioning. Our data suggest that, in spite of normal neuropsychological performance on classic tests, significant differences exist between BD patients and healthy controls in more complex multitasking executive tasks. These tasks seem to pose real-life demands, are supposed to be less strongly associated with fluid intelligence, and are thus potentially more capable of detecting subtle and otherwise unnoticed deficits. The results of this study may indicate that tests of this type can capture subtle deficits that cannot be otherwise detected with classic measures. Minimal changes in executive functioning certainly influence occupational and social functioning. The inclusion of ecological tests in the assessment of BD patients can contribute to providing a more realistic cognitive profile of this patient population. This will undoubtedly allow for a better design of therapeutic and rehabilitation strategies that could help to minimize the impact of cognitive dysfunction on real-life settings and restore psychosocial functioning by means of new ecological cognitive techniques.

Acknowledgement

The present study was funded by a Foundation Institute of Cognitive Neurology (FINECO) grant.

Disclosures

The authors of this paper do not have any commercial associations that might pose a conflict of interest in connection with this manuscript.

References

1. Robinson LJ, Thompson JM, Gallagher P et al. A meta-analysis of cognitive deficits in euthimic patients with bipolar disorder. J Affect Disord 2006; 93: 105–115.
2. Torres IJ, Boudreau VG, Yatham LN. Neuropsychological functioning in euthymic bipolar disorder: a meta-analysis. Acta Psychiatr Scand 2007; 434: 17–26.
3. Zubieta JK, Huguelet P, O'Neil RL, Giordani BJ. Cognitive function in euthymic bipolar I disorder. Psychiatry Res 2001; 102: 9–20.
4. Dickerson F, Boronow JJ, Stallings C, Origoni AE, Cole SK, Yolken RH. Cognitive functioning in schizophrenia and bipolar disorder: comparison of performance on the Repeatable Battery for the Assessment of Neuropsychological Status. Psychiatry Res 2004; 129: 45–53.
5. Martínez-Arán A, Vieta E, Colom F et al. Cognitive dysfunctions in bipolar disorder: evidence of neuropsychological disturbances. Psychother Psychosom 2000; 69: 2–18.
6. Jaeger J, Berns S, Loftus S, Gonzalez C, Czobor P. Neurocognitive test performance predicts functional recovery from acute exacerbation leading to hospitalization in bipolar disorder. Bipolar Disord 2007; 9: 93–102.
7. Tabarés-Seisdedos R, Balanzá-Martínez V, Sánchez-Moreno J et al. Neurocognitive and clinical predictors of functional outcome in patients with schizophrenia and bipolar I disorder at one-year follow-up. J Affect Disord 2008; 109: 286–299.
8. Martino DJ, Marengo E, Igoa A et al. Neurocognitive and symptomatic predictors of functional outcome in bipolar

disorders: a prospective 1 year follow-up study. J Affect Disord 2009; 116: 37–42.
9. Altshuler LL, Ventura J, van Gorp WG, Green MF, Theberge DC, Mintz J. Neurocognitive function in clinically stable men with bipolar I disorder or schizophrenia and normal control subjects. Biol Psychiatry 2004; 56: 560–569.
10. Altshuler L, Bookheimer S, Townsend J et al. Regional brain changes in bipolar I depression: a functional magnetic resonance imaging study. Bipolar Disord 2008; 10: 708–717.
11. Martino DJ, Marengo E, Igoa A et al. Neurocognitive and symptomatic predictors of functional outcome in bipolar disorders: a prospective 1 year follow-up study. J Affect Disord 2008; 116: 37–42.
12. Gioia GA, Isquith PK. Ecological assessment of executive function in traumatic brain injury. Dev Neuropsychol 2004; 25: 135–158.
13. Gregory CA, Hodges JR. Frontotemporal dementia: use of consensus criteria and prevalence of psychiatric features. Neuropsychiatry Neuropsychol Behav Neurol 1996; 9: 145–153.
14. Gregory CA, Lough S, Stone V et al. Theory of Mind in patients with frontal variant frontotemporal dementia and Alzheimer's disease: theoretical and practical implications. Brain 2002; 125: 752–764.
15. Burgess PW, Alderman N, Volle E, Benoit RG, Gilbert SJ. Mesulam's frontal lobe mystery re-examined. Restor Neurol Neurosci 2009; 27: 493–506.
16. Torralva T, Gleichgerrcht E, Lischinsky A, Roca M, Manes F. ''Ecological'' and highly demanding executive tasks detect real life deficits in high functioning adult ADHD patients. J Attentional Disord 2010 (in press).
17. Torralva T, Roca M, Gleichgerrcht E, Bekinschtein T, Manes F. A neuropsychological battery to detect specific executive and social cognitive impairments in early frontotemporal dementia. Brain 2009; 132: 1299–1309.
18. Roca M, Parr A, Thompson R et al. Executive function and fluid intelligence after frontal lobe lesions. Brain 2010; 133: 234–247.
19. Roca M, Torralva T, Meli F et al. Cognitive deficits in multiple sclerosis correlate with changes in fronto-subcortical tracts. Mult Scler 2008; 14: 364–369.
20. Gleichgerrcht E, Torralva T, Roca M, Manes F. Utility of an abbreviated version of the executive and social cognition battery in the detection of executive deficits in early behavioral variant frontotemporal dementia patients. J Int Neuropsychol Soc 2010; 4: 687–694.
21. First M, Spitzer R, Gibbon M, Williams J. Structured Clinical Interview for DSM-IV Axis I Disorders-Clinical Version (SCID-CV). Washington, DC: American Psychiatry Press, 1996.
22. Hamilton A. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960; 23: 56–62.
23. Colom F, Vieta E, Martínez-Arán A et al. Spanish version of a scale for the assessment of mania: validity and reliability of the Young Mania Rating Scale. Med Clin 2002; 28: 119.
24. Beck AT, Steer RA, Brown GK. Manual for the Beck Depression Inventory-II. San Antonio: Psychological Corporation, 1996.
25. Diagnostic and Statistical Manual of Mental Disorders, 4th ed. Washington, DC: American Psychiatric Association, 1994.
26. Burin DI, Jorge RE, Arizaga RA, Paulsen JS. Estimation of premorbid intelligence: the word accentuation test – Buenos Aires version. J Clin Exper Neuropsychol 2000; 22: 677–685.
27. Mathuranath PS, Nestor PJ, Berrios GE, Rakowicz W, Hodges JR. A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. Neurology 2000; 55: 1613–1620.
28. Wechsler D. Wechsler Adult Intelligent Scale-III. San Antonio: The Psychological Corporation, 1997.
29. Partington JE, Leiter RG. Partington's pathway test. Psychol Center Bullet 1949; 1: 9–20.
30. Wechsler D, Stone CP. Wechsler Memory Scale–Revised (WMS-R). San Antonio: The Psychological Corporation, 1987.
31. Rey A. L'examen physiologique dans le cas d'encephalopathie traumatique. Archiv Psychologie 1941; 28: 286–340.
32. Goodglass H, Kaplan E. Boston Diagnostic Aphasia Examination (BDAE). Philadelphia: Lea & Febiger, 1983.
33. Spreen O, Benton AL. The Neurosensory Center Comprehensive Examination for Aphasia. Neuropsychology Laboratory: University of Victoria, 1977.
34. Nelson HE. A modified card sorting test sensitive to frontal lobe deficits. Cortex 1976; 12: 313–324.
35. Dubois B, Slachevsky A, Litvan I, Pillon B. The FAB: Frontal Assessment Battery at bedside. Neurology 2000; 55: 1621–1626.
36. Burgess P. Development of a simplified version of the multiple errands test for use in hospital settings. Neuropsychol Rehabil 2002; 12: 231–255.
37. Manly T, Hawkins K, Evans J, Woldt K, Robertson IH. Rehabilitation of executive function: a facilitation of effective goal management on complex tasks using periodic auditory alerts. Neuropsychologia 2002; 40: 2671–2681.
38. Shallice T, Burgess PW. Deficits in strategy application following frontal lobe damage in man. Brain 1991; 114: 727–741.
39. Krueger CE, Bird AC, Growdon ME, Jang JY, Miller BL, Kramer JH. Conflict monitoring in early frontotemporal dementia. Neurology 2009; 73: 349–355.
40. Gregory CA. Frontal variant of frontotemporal dementia: a cross-sectional and longitudinal study of neuropsychiatric features. Psychol Med 1999; 29: 1205–1217.
41. Goldstein LH, Bernard S, Fenwick PB, Burgess PW, McNeil J. Lateral frontal lobectomy can produce strategy application disorder. J Neurol Neurosurg Psychiatry 1993; 3: 274–276.
42. Martínez-Aran A, Vieta E, Torrent C et al. Functional outcome in bipolar disorder: the role of clinical and cognitive factors. Bipolar Disord 2007; 9: 103–113.
43. O'Shea R, Poz R, Michael A, Berrios GE, Evans JJ, Rubinsztein JS. Ecologically valid cognitive tests and everyday functioning in euthymic bipolar disorder patients. J Affect Disord 2010; 125: 336–340.