

THE FIVE DIGITS TEST ON THE ASSESSMENT OF PSYCHIATRIC PATIENTS
WITH HETEROGENEOUS EDUCATIONAL BACKGROUNDS:
EVIDENCES OF VALIDITY ON THE ASSESSMENT OF BIPOLAR DISORDER

Jonas Jardim de Paula, Suzana Abrantes, Fernando Silva Neves, Leandro Fernandes Malloy-Diniz

Abstract

Objective: to analyze the validity of the Five Digits Test (FDT), an adapted version of the Stroop Color-Word Test (SCWT) on Bipolar Disorder patients (BD).

Method: we assessed 35 BD patients diagnosed in a clinical interview by a psychiatrist using the MINI-Psychiatric Interview. The participants were euthymic during the cognitive assessment. Besides the FDT, the participants performed the SCWT, the Five Points Test (FPT) and the Rey Auditory Verbal Learning Test (RAVLT) as measures of criterion-related, convergent and divergent validity, respectively. We adopted partial correlations, controlling for age, education, sex and symptoms of depression and mania, to analyze the association between the FDT times and errors with the other cognitive measures. Group comparisons with a paired control sample were used to assess the FDT discriminant validity.

Results: we found significant correlations between the FDT and the SCWT and between the FDT and the FPT, but not between the FDT and the RAVLT. The shared variance of the measures sustained the construct validity for the assessment of processing speed and executive functions. Group comparisons indicate large effect sizes for the FDT timed measures and errors in the executive functions components. The test accuracy for group classification was 90%.

Conclusions: the results suggest that the FDT is a valid measure of processing speed and executive functions for the assessment of BD patients, and might be an alternative to the SCWT for patients with low formal education.

Key words: Five Digits Test (FDT), Stroop Color-Word Test (SCWT), Bipolar Disorder patients (BD), MINI-Psychiatric Interview, Rey Auditory Verbal Learning Test (RAVLT)

Declaration of interest: none

Jonas Jardim de Paula^{1,2,3}, Suzana Abrantes⁴, Fernando Silva Neves^{2,5}, Leandro Fernandes Malloy-Diniz^{2,3,5}

¹Departamento de Psicologia, Faculdade de Ciências Médicas de Minas Gerais, Belo Horizonte, Brazil.

²INCT de Medicina Molecular, Faculdade de Medicina, Universidade Federal de Minas Gerais, Av Alfredo Balena, 190, Belo Horizonte-MG, CEP 30130-100, Brazil.

³Laboratório de Investigações em Neurociência Clínica (LINc) – Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.

⁴Private Practice

⁵Departamento de Saúde Mental, Faculdade de Medicina, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.

Corresponding author

Jonas Jardim de Paula

Avenue Alfredo Balena, 190, Belo Horizonte-MG, CEP 30130-100, Brazil. Telephone/Fax: 55-31-9990-2760.

Email: jonasjardim@gmail.com

Introduction

The assessment of processing speed and executive functions in Bipolar Disorder (BD) is an important aspect on their clinical evaluation. Deficits in selective attention (a component of one Executive Function called “Inhibitory Control”) is related to several clinical outcomes including impairments in everyday life activities (O’Shea et al. 2010), psychosocial functioning (Solé et al. 2012), and suicide behavior (Malloy-Diniz et al. 2011). One of the most used paradigms for this cognitive function is the interference control, assessed by Stroop-like tests, as the Stroop Color-Word Test (SCWT). It involves the time and errors in colour naming,

naming the colours of graphed words and naming the colours of graphed colour-names in incongruent (Red-Green) conditions. The first two components are usually related to processing speed while the last is related to executive functions (Troyer et al. 2006).

The SCWT has several limitations in the assessment of patients with lower formal education, learning disabilities or non-clinical difficulties in word reading (Cox et al. 1997). The proficiency in reading is an important factor for the interpretation of the SCWT. In Brazil, the educational background of most adults, especially those in public health care units is often deficient, with difficulty in reading fluency and comprehension of small texts or even isolated words.

An alternative to the classical SCWT is the numeric or counting versions of the test. These are Mini-Verbal or Language-Free versions of the SCWR: they use a very simple content (number, colours or figures) but the same design as classical tests do (Sedó and DeCristoforo, 2001).

The Five Digits Test (FDT) is a numeric-Stroop test, which is relatively independent of reading routines (Sedó 2007). The test involves four different components. On the first one, the patient must read one out of five numbers in a congruent condition (“1”, he/she must say “one”; or “3-3-3”, he/she must say three), exposed on small squares (FDT Reading). On the second component, he/she must tell how many asterisks (1 to 5) are in each square (“*”, he/she must say “one”; or “***” he/she must say “three”) (FDT Counting). The third component involves an interference condition where one to five numbers are exposed in an incongruent condition (“2-2-2” or “5-5-5-5”), and the subject must not say the name of the number, but tell how many numbers are in each stimulus (in the example, for “2-2-2” he/she must say “three” and for “5-5-5-5” he/she must say “four”) (FDT Choosing). The last component is similar to the third one, but for every five stimuli there is one marked square where he/she must not say how many numbers are in each square, but read one of the numbers (FDT Shifting). The time and errors of each of the four conditions are used as measures of “automatic attentional processes” (Reading and Counting) and “controlled attentional processes” (Choosing and Shifting). The two first components are more related to simple attention and processing speed, while the two last are more dependent on selective attention and executive functions.

Although the test is a promising measure of these constructs for the Brazilian population, we found no studies assessing its psychometric properties in psychiatric patients. In this brief report, we aim to assess the FDT construct, criterion-related and discriminant validity in BD patients.

Methods

Participants

We assessed 35 euthymic BD patients from a specialized health unit of mental health in a university hospital. We performed the diagnosis based on the DSM-IV-TR criteria by a psychiatric interview conducted by a certified clinical psychiatrist. On this assessment, the professional also applied the Portuguese version of the MINT International Neuropsychiatric Interview (Amorim 2000). The mean age of the patients was 46.21 (\pm 13.12) years and 71% had eleven or less years of formal education. The BD patients had been euthymic for at least four weeks according to the clinical interview. The mean Beck Depression Inventory (Beck et al. 1961) score was 8.83 (\pm 2.44) and the Young Mania Rating Scale (Young et al. 1978) was 6.26 (\pm 2.38), scores not suggestive of prominent symptoms of depression or mania during the neuropsychological assessment, characterizing the euthymic phase of BD.

Neuropsychological assessment

We used four different neuropsychological measures. Our interest was to assess the FDT construct and criterion-related validity, so we chose a criterion

measure, a convergent measure (another test of processing speed and executive functions) and a divergent measure (a test of another cognitive domain).

Criterion-Related validity: For this purpose, we selected the SCWT Victoria version (Troyer et al. 2006). The test has three components: Box, Neutral Words and Colour Words. The executive demands of the tasks increase along their three parts. We used the SCWT times to assess the criterion-related validity of the FDT timed measures and the total errors of each SCWT component for the FDT error measures.

Convergent validity: We used the Five Points Test (Regard et al. 1982) for the assessment of this psychometric property. The Five Points Test is a non-verbal fluency task where the patient must draw a series of figures using straight lines in several squares containing five dots. The test demands, besides the motor fluency, planning, monitoring and cognitive shifting, in order to produce new drawing but avoid errors or repetitions. We selected two measures from this test: the total of correct drawings as a convergent measure for the FDT times (a more pure measure of fluency and processing speed) and the total number of errors as a convergent measure for the FDT errors (a more pure measure of executive functions, including inhibitory control and cognitive shifting).

Divergent Validity: For the assessment of this psychometric property, we selected the Brazilian Portuguese version of the Rey Auditory-Verbal Learning Test (Malloy-Diniz et al. 2007). This test is a classical measure of episodic memory, validated for the Brazilian adult population. As a general estimate of episodic memory, we selected the sum of words (RAVLT Total) of the five learning trials.

Discriminant validity: For the assessment of this psychometric property, we assessed a paired (1:1) control group without evidences of cognitive, psychiatric or neurological disorders. The control group was a convenience sample of healthy adults matched to the BD patients in the distribution of sex ($\chi^2=0.08$, $p=0.771$), age ($\chi^2=0.56$ $p=0.757$) and formal education ($\chi^2=0.58$, $p=0.445$). The control group did not report psychiatric disorders or neurological diseases and had scores on the Standard Raven Progressive Matrices Test (Raven, 1981) above the percentile 10 according to the Brazilian population norms.

Statistical analysis

We explored the relationship of the FDT measures with the other tests computing partial correlations of the neuropsychological measures controlling for the effects of age, education and symptoms of depression and mania. We selected this procedure since sociodemographic variables may influence neuropsychological test performance and might inflate the correlations of the cognitive constructs. Measures that do not follow a normal distribution underwent logarithmic, square root, square or cubic transformations to better fit the analysis assumptions. Since multiple correlations were computed and the sample size was relatively small, a more conservative p-value was selected to avoid type-I errors ($p=0.01$). The BD and control groups were compared by independent samples t-tests, and the Cohen's d was adopted as a measure of effect sizes. We also performed a stepwise binary logistic regression model to analyze the FDT accuracy in the group classification. The test times and errors were entered as independent predictors.

Results

The BD participants description is shown on **table 1**. **Table 2** shows the correlations of the neuropsychological measures divided by the validity assessed (criterion-related, convergent and divergent). On the criterion-related validity analysis, we found significant correlations between the FDT automatic components times (Reading and Counting) with all the timed measures of the SCWT, ranging from 22% to 57% of shared variance. The correlations of all FDT

timed measures with the SCWT Color-Words time were significant, and the shared variance ranged from 33% to 57%. The pattern of errors was more specific, with significant correlations between the FDT Reading errors only with SCWT Dots and Neutral Words (30 and 38% of shared variance, respectively). FDT Counting errors had no correlation with the criterion measure. The executive components' errors of the FDT (Choosing and Shifting) were only associated with the SCWT Color-Word component (69% and 49% of shared variance, respectively). For the convergent and divergent validity

Table 1. Bipolar Disorder patient's description

Participants description		% or M(SD)
Gender	Male	20%
	Female	80%
Age Group	18 to 35 years	29%
	36 to 50 years	26%
	51 to 65 years	46%
Formal Education	Less than 12	71%
	12 years or more	29%
Marital Status	Single	51%
	Married	23%
	Widowed	9%
	Divorced	9%
	Non-Informed	9%
Beck Depression Inventory		8.83 (2.44)
Young Mania Rating Scale		6.26 (2.38)
SCWT 1 (Time)		20.38 (5.67)
SCWT 2 (Time)		26.62 (10.76)
SCWT 3 (Times)		39.74 (18.66)
SCWT - Interference Score		16.24 (13.97)
SCWT 1 (Errors)		0.21 (0.73)
SCWT 2 (Errors)		0.15 (0.70)
SCWT 3 (Errors)		1.94 (4.68)
Five Points Test (Unique Drawings)		19.63 (2.21)
Five Points Test (Errors)		0.87 (2.21)
RAVLT (Sum of Words)		40.94 (9.79)

M: Mean, SD: Standard-Deviation, SCWT: Stroop Color-Word Test, RAVLT: Rey Auditory-Verbal Learning Test

Table 2. Partial correlations and shared variance (r^2) between the Five Digits Test with the Stroop Color-Word, Five Points Test and Rey Auditory Verbal-Learning Tests

FDT Times	SCWT Dots	SCWT Neutral-W	SCWT Color-W	Five Points Test	RAVLT
Reading (T) ¹	0.579* (34%)	0.659* (43%)	0.754* (57%)	-0.471* (22%)	-0.164 (3%)
Counting (T) ¹	0.554* (31%)	0.715* (51%)	0.612* (37%)	-0.440* (19%)	-0.047 (0%)
Choosing (T) ¹	0.394 (16%)	0.474* (22%)	0.573* (33%)	-0.238 (6%)	-0.204 (4%)
Shifting (T) ¹	0.393 (15%)	0.476* (23%)	0.528* (28%)	-0.177 (3%)	-0.082 (1%)
Reading (E) ²	0.550* (30%)	0.618* (38%)	0.38 (14%)	0.508* (26%)	0.034 (0%)
Counting (E) ²	0.375 (14%)	0.408 (17%)	0.315 (10%)	0.491* (24%)	-0.121 (1%)
Choosing (E) ²	0.117 (1%)	0.559* (31%)	0.831* (69%)	0.746* (56%)	0.114 (1%)
Shifting (E) ²	0.119 (1%)	0.146 (2%)	0.698* (49%)	0.526* (28%)	0.097 (1%)

Note: 1 – Correlations with SCWT times and Five Points Test unique drawings. 2- Correlations with SCWT errors and Five Points Test errors. * p<0.01

FDT: Five Digits Test, T: Times, E: Errors, SCWT: Stroop Color Word Test, W? words, RAVLT: Rey Auditory-Verbal Learning Test sum of words

we found correlations between the Five Points Test unique drawing (a more “pure” measure of fluency and processing speed) with the two first FDT components (22 and 18% of shared variance respectively), but not with the Choosing and Shifting components. However, the Five Points Test errors were significantly correlated with all the FDT errors, but more with the Choosing Errors (56% of shared variance). On the divergent validity analysis, no significant correlations were found between any measures of the FDT with the RAVLT sum of words.

Table 3 shows the BD and control groups performance in the FDT, as well as the group comparisons. We found large effect sizes for all FDT timed measures (Cohens’s *d* ranging from 2.8 to 3.3) and for errors of Choosing (*d*=1.1) and Shifting (*d*=1.0). Comparisons on “automatic” attentional process errors were not significant. When all the timed measures along with the Choosing and Shifting Errors were entered on a forward stepwise binary logistic regression model, the procedure yielded significant results on its last step ($\chi^2=62.81$, *df*=1, *p*<0.001), and answered for 79% of variance according to the Nagelkerke pseudo-*R*² test. The FDT Shifting time alone ($\beta= -0.32$, *SE*=0.09, *p*<0.001, *Wald*=12.43) had an overall classification rate of 90%, 86% for the BD group and 94% for the control group. The results suggest a sensitivity of 0.94, specificity of 0.87, predictive positive value of 0.86 and predictive negative value of 0.94 for the FDT Shifting time in the distinction between controls and BD patients.

Discussion

Our results provide a series of evidences about the construct validity of the FDT for the assessment of patients with BD. When correlated with a criterion measure of selective-attention/executive functions (SCWT) we found a large fraction of shared variance between the tests, both in terms of time and errors. When correlated to the convergent measure, a non-verbal fluency test (Five Points Test), we found significant associations between the automatic timed components of the test and the production of drawings; and the errors on the FDT with incorrect drawings

on the convergent task. Finally, we found a divergent pattern of association between the FDT and the RAVLT, an evidence of divergent validity. Significant group differences were found for the FDT times (with very large effect sizes), for FDT Choosing and Shifting Errors (with large effect sizes), but not for Reading and Counting errors. In a binary logistic regression model, the FDT Shifting time predicted group membership with high accuracy (90%).

The results are convergent with other research findings, which examined the association of the FDT measures with other cognitive tests. Sedó and DeCristoforo (2001) revised correlational studies of the FDT with other measures and found a significant association with the SCWT. The effect sizes reported by the authors were very similar to ours, suggesting a transcultural validity. Sedó (2007) also reviewed other correlational studies and reported significant correlations with other attentional and executive functions tests with moderate or high effect sizes. In the Brazilian population, Magalhães (2013) examined the association of the FDT with other measures related to inhibitory control in healthy adults. Her results indicated that the FDT along with the SCWT and other tests of attention composed a latent factor related to interference control, a subcomponent of the inhibitory control system. A study with mood disorders carried in older adults with major depressive disorders also revealed significant group differences between patients and controls, with moderate to high effect sizes (de Paula et al. 2011).

This study has important limitations, which we must address. We had a small sample size. Although the participants were well characterized and some constraint were adopted for the statical analysis, the sample size is not sufficient to identify associations with small effect sizes. In addition, other factors, which may have influenced the results, such as premorbid IQ, number of BD episodes, reading and counting proficiency, were not controlled. Despite these limitations, the results show evidence of validity for a new attention/executive functions test, which may be more adequate for the assessment of Brazilian adults with psychiatric disorders than the traditional ones due to their heterogeneous educational background.

Table 3. Group comparisons on the FDT measures between the Bipolar Disorder group and a paired control group (matched by gender, age and education)

FDT Measures	Bipolar Disorder (n=35)		Controls (n=35)		Independent samples t-test	
	M (SD)	Min-Max	M (SD)	Min-Max	t	d
Reading (T)	29.83 (8.50)	16-55	19.94 (2.59)	14-24	6.6*	3.3
Counting (T)	37.91 (12.27)	19-73	23.43 (2.94)	19-29	6.8*	3.1
Choosing (T)	57.94 (21.17)	31-120	35.06 (5.09)	26-46	6.2*	2.8
Shifting (T)	79.71 (29.98)	41-153	44.17 (5.53)	35-55	6.9*	2.9
Reading (E)	0.29 (0.96)	0-4	0.00 (0.00)	0-0	1.8	-
Counting (E)	0.49 (1.54)	0-8	0.00 (0.00)	0-0	1.9	-
Choosing (E)	1.97 (3.71)	0-17	0.29 (0.57)	0-2	2.7*	1.1
Shifting (E)	3.09 (5.35)	0-28	0.97 (1.12)	0-3	2.3*	1.0

Note: **p*<0.05

T: Time, E: Errors, M: Mean, SD: Standard-Deviation, Min: Minimum, Max: Maximun, , d: Cohen’s *d* (effect size)

Acknowledgements

Grant INCT-MM (FAPEMIG: CBB-APQ-00075-09/CNPq 573646/2008-2).

References

- Amorim P (2000). Mini International Neuropsychiatric Interview (MINI): validation of a short structured diagnostic psychiatric interview. *Revista Brasileira de Psiquiatria* 22, 3, 106-115.
- Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J (1961). An inventory for measuring depression. *Archives of General Psychiatry* 4, 561-571.
- Cox CS, Chee E, Chase GA, Baumgardner L, Schuerholz LJ, Reader MJ, Morh J, Denckla MB (1997). Reading proficiency affects the construct validity of the Stroop Test Interference Score. *The Clinical Neuropsychologist* 11, 2, 105-110.
- de Paula JJ, Ávila RT, Costa DS, Moraes EN, Bicalho MA, Nicolato R, Corrêa H, Sedó M, Malloy-Diniz LF (2011). Assessing processing speed and executive functions in low educated older adults: the use of the Five Digits Test in patients with Alzheimer's disease, mild cognitive impairment and major depressive disorders. *Clinical Neuropsychiatry: Journal of Treatment Evaluation* 8, 6, 339-346.
- Magalhães SS (2013). *Estrutura fatorial do controle inibitório no envelhecimento: Comparação entre amostras de adultos e idosos* [dissertation]. Curitiba: Universidade Federal do Paraná.
- Malloy-Diniz LF, Neves FS, Abrantes SS, Fuentes D, Corrêa H (2009). Suicide behavior and neuropsychological assessment of type I bipolar patients. *Journal of Affective Disorders* 112, 1, 231-236.
- O'Shea R, Poz R, Michael A, Berrios GE, Evans JJ, Rubinsztein JS (2010). Ecologically valid cognitive tests and everyday functioning in euthymic bipolar disorder patients. *Journal of Affective Disorders* 125, 1, 336-340.
- Raven J (1981). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Harcourt Assessment, San Antonio, TX.
- Regard M, Strauss E, Knapp P (1982). Children's production on verbal and non-verbal fluency tasks. *Perceptual and Motor Skills* 55, 1, 839-44.
- Malloy-Diniz LF, Lasmar VA, Gazinelli LS, Fuentes D, Salgado JV (2007). The Rey Auditory-Verbal Learning Test: applicability for the Brazilian elderly population. *Revista Brasileira de Psiquiatria* 29, 4, 324-329.
- Sedó MA (2007). *Five Digits Test*. TEA Ediciones, Madrid.
- Sedó MA, DeCristoforo L (2001). All-language verbal tests free from linguistic barriers. *Revista Espanola de Neurociología* 3, 3, 68-82.
- Solé B, Bonnin CM, Balanzá-Martínez V, Tabarés-Seisdedos R, Popovic D, Martínez-Arán A, Vieta E (2012). Neurocognitive impairment and psychosocial functioning in bipolar II disorder. *Acta Psychiatrica Scandinavica* 125, 4, 309-317.
- Troyer AK, Leach L, Strauss E (2006). Aging and response inhibition: Normative data for the Victoria Stroop Test. *Aging, Neuropsychology, and Cognition: A Journal on Normal and Dysfunctional Development* 13, 1, 20-35.
- Young RC, Biggs JT, Ziegler VE, Meyer DA (1978). A rating scale for mania: reliability, validity and sensitivity. *The British Journal of Psychiatry* 133, 429-435.