

Clinical Spanish Norms of the Stroop Test for Traumatic Brain Injury and Schizophrenia

Genny Lubrini¹, José A. Periañez², Marco Rios-Lago³, Raquel Viejo-Sobera⁴, Rosa Ayesa-Arriola⁵, Ignacio Sanchez-Cubillo⁶, Benedicto Crespo-Facorro⁵, Juan Álvarez-Linera⁷, Daniel Adrover-Roig⁸ and José M. Rodriguez-Sanchez⁵

¹ Hospital Universitario La Paz (Spain)

² Universidad Complutense (Spain)

³ UNED (Spain)

⁴ Universitat Oberta de Catalunya (Spain)

⁵ CIBERSAM, Centro Investigación Biomédica en Red Salud Mental (Spain)

⁶ Red Menni de Servicios de Atención al Daño Cerebral (Spain)

⁷ Hospital Ruber Internacional, Madrid (Spain)

⁸ Universidad de las Islas Baleares (Spain)

Abstract. The Stroop Color-Word Test is a useful tool to evaluate executive attention and speed of processing. Recent studies have provided norms for different populations of healthy individuals to avoid misinterpretation of scores due to demographic and cultural differences. In addition, clinical norms may improve the assessment of cognitive dysfunction severity and its clinical course. Spanish normative data are provided for 158 closed traumatic brain injury (TBI) and 149 first-episode schizophrenia spectrum disorder (SCH) patients. A group of 285 Spanish healthy individuals (HC) was also considered for comparison purposes. Differences between groups were found in all Stroop scores with HC outperforming both clinical groups ($p < .002$ in all cases; $d > .3$ in all cases). TBI patients scored lower than SCH patients in word-reading ($p < .001$ and $d = .6$), and color-naming conditions ($p < .001$ and $d = .4$), but not in the color-word condition ($p = .34$ and $d = .03$). However, SCH patients exhibited a higher interference effect as compared to TBI ($p < .002$ and $d = .5$). Three sets of norms stratified by age and education (HC), and by education (TBI and SCH) are presented for clinical use.

Received 20 July 2014; Revised 26 September 2014; Accepted 7 October 2014

Keywords: attention, clinical norms, interference control, schizophrenia, traumatic brain injury.

The Stroop Color-Word Test is a commonly used tool in clinical and experimental psychological settings as a measure of selective attention, cognitive flexibility, and processing speed (Strauss, Sherman, & Spreen, 2006). Stroop paradigms include an interference task where the subject has to inhibit a highly automatic response in benefit of a less automatic one (Jensen & Rohwer, 1966). This is generally achieved asking the subject to name the ink color of a word whose

meaning is incongruent with it (e.g. to name the red ink color of the word “blue”; Jensen & Rohwer, 1966). Under these conditions a “color-word interference effect” emerges as a significant increase in the time required completing the incongruent task as compared to the control task (i.e., word reading or congruent color naming). While word reading and color naming conditions have been proposed as measures of processing speed, the interference effect is intended to provide a measure of executive attention (Ríos, Periañez, & Muñoz-Céspedes, 2004; Strauss et al., 2006).

The generalized use of Stroop test versions in psychological assessment such as the one by Golden (1978) has boosted the publication of different norms stratified by a number of demographic variables. Thus, age, educational level, or sex have been suggested as relevant features modulating task performance (Moering, Schinka, Mortimer, & Graves, 2004; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2006). In addition, some studies have demonstrated that the specific weight of demographic variables in Stroop scores may vary when comparing different ethnic groups,

Correspondence concerning this article should be addressed to Jose A. Periañez. Dept. Psicología Básica II. Facultad de Psicología. Universidad Complutense de Madrid. Campus de Somosaguas. 28223. Madrid (Spain). Phone: +34-913943181. Fax: +34-913943189.

E-mail: japerianez@psi.ucm.es

This work was partially supported by the grant PSI2009-14415-C03-03 from the Ministerio de Ciencia e Innovación (MICINN) of Spain; by MAPFRE Medicina Foundation; Instituto de Salud Carlos III PI020499, PI050427, PI060507; Plan Nacional de Drogas Research Grant 2005- Ordensco/3246/2004; SENY Foundation Research Grant CI 2005-0308007; and Fundación Marqués de Valdecilla API07/011. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

We want to thank all patients and healthy participants who voluntarily and generously took part in the study.

even within the same country, thus justifying specific norms or demographic corrections for them (Norman et al., 2011). Following this rationale, recent works have provided Stroop norms for large samples of Dutch (Van der Elst et al., 2006), Greek (Zalonis et al., 2009), Korean (Seo et al., 2008), African and Caucasian North-America (Norman et al., 2011), and Spanish populations (Peña-Casanova et al., 2009; Rognoni et al., 2013).

While the primary function of norms is to identify the presence of pathological performance, they have been secondarily used to scale cognitive impairment. At this regard, norms from healthy samples are generally sensitive to achieve the first aim. However, it has been shown that they may present serious limitations regarding the second one, i.e., a large percentage of patients often score out of the range established by healthy norms. For instance, in a previous normative study of another attentional test (the Trail Making Test; Periañez et al., 2007) a 30 % of the schizophrenia patients, and a 70 % of the traumatic brain injury (TBI) patients from the high education groups scored below percentile 5 as established by healthy norms (see Tables 6, 7, and 8 in Periañez et al., 2007). This fact highlights at least two main problems when assessing clinical performance using norms from healthy samples. Firstly, norms from healthy samples cannot accurately scale cognitive impairment. Secondly, they lack of sensitivity for detecting subtle clinical changes across time (even after correcting for learning effects in retest measures; Van der Elst, Molenberghs, Van Boxtel, & Jolles, 2013). These are two central concerns in clinical settings for both prognostic purposes and assessment of patients' clinical course (Kizilbash, Warschawsky, & Donders, 2001; Periañez et al., 2007). For these reasons, different works have provided clinical norms for different tests such as Wisconsin Card Sorting Tests (Iverson, Slick, & Franzen, 2000), Trail Making Test (Periañez et al., 2007), Rey's Auditory Verbal Learning Test (Badcock, Dragovic, Dawson, & Jones, 2011), or Boston Naming Test (Casals-Coll et al., 2014). However, to our knowledge, no clinical norms exist for the Stroop test.

Impaired performance on the Stroop has been described in a wide variety of clinical groups characterized by executive control and prefrontal lobe dysfunction such as traumatic brain injury (Felmingham, Baguley, & Green, 2004; Ríos et al., 2004), schizophrenia (Hepp, Maier, Hermle, & Spitzer, 1996; Rodríguez-Sánchez et al., 2007), or even during normal ageing (Coubard et al., 2011; Mayas, Fuentes, & Ballesteros, 2012). Among others, these groups may constitute target populations for the design of specific clinical norms. In spite of this, until now, most normative studies on the Stroop test have focused in non-clinical samples.

The main purpose of the present study was to provide clinical norms of Golden's version of the Stroop test (Golden, 1978; 1999) for schizophrenia and TBI Spanish populations together with matching data for healthy individuals. These norms will offer the clinician a tool for a comprehensive description of patients according to severity, as well as a more sensitive measure of clinical course.

Method

Participants

A total of 592 subjects took part in the study: 285 healthy subjects (170 female); 158 closed traumatic brain injury patients (31 female). 149 first-episode schizophrenia patients (67 female). They all were Spanish speakers and had normal or corrected-to-normal vision.

Healthy Controls (HC) were recruited from undergraduate university classes, university staff, social organizations, hospitals, and health care centers from three different regions of Spain (Madrid, Bilbao, and Santander). Medical complications, psychiatric disturbance, substance abuse (excluding nicotine), or neurological disease diagnosis were criteria for exclusion in this sample.

The schizophrenia (SCH) sample was comprised of patients with diagnosis of schizophrenia spectrum disorders (schizophrenia, schizophreniform, or schizoaffective) in their first episode. All patients from this group attended at a program for first-episode psychosis (PAFIP) carried out at the Hospital Marques de Valdecilla in Santander (see a detailed description in Crespo-Facorro et al., 2006). Diagnoses were confirmed by an experienced psychiatrist by means of the Structured Clinical Interview for DSM-IV (SCID-I) 6 months after the initial contact. Patients with mental retardation, neurological illness, or drug dependence (excluding nicotine) were excluded. None of the patients had received neuroleptic medication prior to contact with the program. However, they all were on neuroleptic medication and had reached clinical stabilization when neuropsychological assessment for the current study was conducted. Cognitive assessment was therefore performed at 10 weeks after pharmacological treatment initiation, which has been previously stated as the most appropriate for cognitive evaluation (González-Blanch et al., 2006).

One hundred and fifty-eight moderate to severe closed traumatic brain injury (TBI) patients were recruited from the Brain Damage Unit at Beata María Ana Hospital in Madrid, and the Brain Damage Unit at Aita Menni Hospital in Bilbao. Glasgow Coma Scale (GCS; Teasdale & Jennett, 1974) was available for 139 subjects (mean 6.9 ± 3.6). Post-traumatic

amnesia duration (PTA), assessed with the Galveston Orientation and Amnesia Test (GOAT; Levin, O'Donnell, & Grossman, 1979) was also recorded in 110 patients (mean 41.8 ± 35.8 days). The mean time since injury was 505 ± 692 days. All TBI patients had at least one of the two scores (GCS and/or GOAT) recorded. Patients with impairments that may interfere with testing (visual difficulties, aphasia, or apraxia) were excluded from this sample. All patients were out of PTA at the time of testing. The specific site of lesion was not considered for the analyses given that closed TBI is generally characterized by a diffuse pattern of injuries difficult to identify using conventional neuroimaging methods (Arfanakis et al., 2002). None of the participants was involved in litigation regarding their injury at the time of testing. All participants were informed about the investigation prior to the psychological evaluation session and signed a consent form according to the Declaration of Helsinki.

Procedures

Participants were administered the Spanish adaptation of the Stroop test (Golden, 1999) by expertise psychologists as a part of a larger battery. This version consists of three conditions: a word-reading condition (WR) with 100 color words in Spanish printed in black ink, a color-naming condition (CN) with 100 "Xs" printed in color (red, green, and blue), and a color-word condition (CW) with the 100 color words in Spanish from the first condition printed in incongruent colors. Subjects were asked to read down the columns starting by the top word in the leftmost column. After 45 seconds, the item last named on each condition was noted. Whenever an error occurred participants were instructed to correct it but time counting did not stop. Direct scores were measured as the number of items completed on each condition. An Interference score (IS) was also calculated as the difference between CW score and CW', where CW' equals $(WR \times CN) / (WR + CN)$. This formula is based on the assumption that the time to name a color-word item is equal to the time needed to suppress the reading of a word plus the time to identify a color (Golden & Freshwater, 2002). In case of impaired interference control, reading the word in the color-word condition will actively interfere with naming the color and switching from one to the other will go slowly, resulting in a smaller color-word score relative to the predicted score and thus a negative Golden's interference score (Lansbergen, Kenemans, & van Engeland, 2007).

Statistical analysis

Chi-square statistic was used to explore sex distribution among groups. One-way analysis of variance

(ANOVA) was used to compare groups regarding age and years of education. A series of analyses of covariance (ANCOVA) were used to explore the presence of differences in Stroop scores among samples, thus justifying the need of different norms for each of them. Correlation analyses and contingency coefficients were carried out for continuous and dichotomous variables, respectively, to explore the most appropriate variables for stratification. In a further step, multiple regression analyses were used on each separate group to study the relative impact of each stratification variable on Stroop scores. Cut-off points for age and education were established using the percentile 50. Finally, independent sample t-tests were performed to assess the appropriateness of cut-off points for each sample. Scores from the resulting groups were transformed into percentile scores. A significance level of .05 was set for all contrasts. A Bonferroni-corrected significance level of $p < .05$ was adopted for all tests of simple effects involving multiple comparisons. Effects sizes (Cohens' d) for all contrasts were calculated with G*Power 3.1 statistical software (Faul, Erdfelder, Lang, & Buchner, 2007).

Results

Demographics

Descriptive statistics for age, education, sex, and Stroop scores are presented in Table 1. The three groups differed on sex distribution ($\chi^2 = 65.7; p < .001; d = .65$). They also differed in age, $F(2, 589) = 15.85; p < .001; d = .24$, and years of education, $F(2, 589) = 21.27; p < .001; d = .27$. *Post hoc* analyses revealed that healthy controls were older than TBI and SCH patients ($p < .001$), but SCH and TBI patients did not differ in age from each other. SCH patients had a lower educational level than both HC and TBI patients ($ps < .001$), being the difference between the latter groups not significant.

Between-Group comparisons

Univariate ANCOVAs were performed in order to address between-group differences in Stroop scores using age and education as covariates to control their influence in performance. There was a main effect of Group in WR, $F(2, 587) = 113.3; p < .001; d = .58$; in CN, $F(2, 587) = 126.7; p < .001; d = .6$; in CW, $F(2, 587) = 133.4; p < .001; d = .57$; and IS, $F(2, 587) = 39.9; p < .001; d = .33$. *Post hoc* analyses revealed that for WR and CN scores, all groups differed from each other ($ps < .001; d > .4$ in all cases). In both cases, HC outperformed both clinical groups, with TBI patients showing the worst performance (see Table 1). For CW score, HC participants outperformed both clinical groups ($ps < .001; d > 1$ in both cases), but clinical groups did not differ from each

Table 1. Statistical properties of demographic and Stroop variables for each sample

	HC		SCH		TBI	
	<i>M (SD)</i>	min-max	<i>M (SD)</i>	min-max	<i>M (SD)</i>	min-max
N	285		149		158	
Sex (m/f)	(115/170)		(82/67)		(127/31)	
Age	39.1 (18.4)	15–80	31.77 (10.6)	16–60	32.23 (12.4)	15–72
Edu	12.6 (3.7)	2–24	10.33 (3.5)	6–17	12.62 (3.7)	7–20
WR	104.4 (16.8)	53–147	93.01 (17.5)	45–140	79.23 (22.8)	8–130
CN	74.97 (15)	20–113	61.43 (12.6)	28–117	56 (16.5)	11–100
CW	47.92 (13.6)	16–84	34.47 (10.3)	7–76	34.16 (13.5)	1–85
IS	4.6 (10)	–25.9–28.81	–2.26 (7.1)	–22.6–17.1	1.66 (9.6)	–30.5–54

Note: HC = Healthy Controls; SCH = Schizophrenia patients; TBI = Traumatic Brain Injury patients; Edu = Education in years; WR = Word Reading; CN = Color Naming; CW = Color Word; IS = Interference Score.

other ($p = .34$; $d = .03$). Finally, all groups differed from each other in the IS ($ps < .002$; $d > .3$ in all cases), with HC participants showing the lowest interference effect, and SCH the highest.

Correlation and Regression analyses

Correlation analyses and contingency coefficients provided an approach to study relationships between Stroop scores and demographic variables (Table 2).

For HC participants, age was negatively correlated with all Stroop scores ($ps < .001$), whereas education correlated positively with WR, CN and CW (see Table 2). For SCH participants education correlated with all Stroop scores but IS, and age correlated negatively with CW and IS. Finally, for the TBI group education correlated positively with all Stroop scores but IS, whereas age correlated negatively with CW and IS. The analyses of contingency coefficients revealed that sex was not significantly related to any Stroop score in any of the three groups ($ps > .508$ for HC; $ps > .199$ for SCH; $ps > .221$ for TBI). Accordingly, only age and education were considered for further analyses.

Multiple regression analyses were carried out to study the relative contribution of age and education to Stroop scores on each sample separately (see Table 3). For the HC sample, age and education considered together had a significant contribution to the prediction of all Stroop variables ($ps < .001$). They jointly accounted for 16.7% of variance of WR score ($d = .18$), 25.9% of CN score ($d = .26$), 38% of CW score ($d = .38$), and 22.4% of IS score ($d = .23$). Partial correlations from multiple regression suggested that age and education were the most appropriate stratification variables for HC. In fact, age accounted for a significant portion of variance of all Stroop variables followed by education that accounted for a significant portion of variance of WR,

Table 2. Correlation matrixes of demographic and Stroop variables for each sample

HC						
	Edu	Age	WR	CN	CW	IS
Edu	1					
Age	-.06	1				
WR	.3**	-.33**	1			
CN	.22**	-.49**	.61**	1		
CW	.22**	-.61**	.46**	.73**	1	
IS	-.09	-.47**	-.01	.03**	.84**	1
SCH						
	Edu	Age	WR	CN	CW	IS
Edu	1					
Age	.23**	1				
WR	.31**	-.04	1			
CN	.22**	-.08	.62**	1		
CW	.29**	-.18*	.56**	.73**	1	
IS	.16	-.2*	.01	.17*	.76**	1
TBI						
	Edu	Age	WR	CN	CW	IS
Edu	1					
Age	.36**	1				
WR	.29**	-.01	1			
CN	.22**	-.08	.79**	1		
CW	.24**	-.16*	.62**	.71**	1	
IS	.09	-.17*	-.02	.06	.73**	1

Note: HC = Healthy Controls; SCH = Schizophrenia patients; TBI = Traumatic Brain Injury patients; Edu = Education in years; WR = Word Reading; CN = Color Naming; CW = Color Word; IS = Interference Score. * $p < .05$; ** $p < .01$ (Two-tailed).

Table 3. Results from multiple regression analyses using Stroop scores as criterion and education and age as predictors

	WR			CN			CW			IS		
	B	<i>p</i>	Var	B	<i>p</i>	Var	B	<i>p</i>	Var	B	<i>p</i>	Var
HC												
Edu	.24	.001	6.7	.14	.01	2.3	.11	.02	1.9	-.01	.89	.01
Age	-.29	.001	9	-.47	.001	22	-.58	.001	35	-.47	.001	22.2
SCH												
Edu	.33	.001	10.6	.25	.002	6.2	.35	.001	12	.22	.009	4.5
Age	-.12	.15	1.4	-.14	.08	2	-.27	.001	7.4	-.25	.002	6.2
TBI												
Edu	.33	.001	9.8	.28	.001	7.2	.35	.001	10.8	.18	.037	2.8
Age	-.13	.114	1.6	-.18	.029	3	-.29	.001	7.8	-.23	.006	4.7

Note: HC = Healthy Controls; SCH = Schizophrenia patients; TBI = Traumatic Brain Injury patients; Edu = Education in years; WR = Word Reading; CN = Color Naming; CW = Color Word; IS = Interference Score; Var = Percent of explained variance.

CN, and CW (see Table 3). Given the large number of participants in this sample, both variables were considered for stratification.

For the SCH sample, age and education considered together had a significant contribution to the prediction of all Stroop variables ($ps < .006$). They jointly accounted for 10.7% of variance of WR ($d = .12$), 7% of CN ($d = .08$), 15% of CW ($d = .16$), and 8.5% of the IS score ($d = .1$). Examination of partial correlations from multiple regression revealed that education accounted for a significant portion of variance of all Stroop scores, while age only accounted for CW and IS scores (see Table 3). For this reason education was the variable considered for stratification.

For the TBI sample, age and education considered together had a significant contribution to the prediction of all Stroop variables ($ps < .012$). They jointly accounted for 10% of variance of WR score ($d = .11$), 8% of CN score ($d = .09$), 13.2% of CW score ($d = .14$), and 4.3% of IS score ($d = .07$). Examination of partial correlations from multiple regressions revealed a significant contribution of both age and education to three of the Stroop variables (see Table 3). However, only education was selected for stratification given its larger contribution to the explanation of the variance as compared to age.

Stratification

First, the HC sample was divided into two age groups according to percentile 50: young and middle-age (see Table 4 for descriptive statistics). Both groups differed in WR, $t(283) = 4.8$; $p < .001$; $d = .58$; in CN score, $t(283) = 8$; $p < .001$; $d = .98$; in CW score, $t(283) = 11.3$; $p < .001$; $d = 1.44$; and IS score, $t(283) = 8.8$; $p < .001$; $d = 1.15$.

Second, each age group was divided into two education groups: young-high education, young-low education, middle age-high education, and middle age-low education. In the young group, high education participants outperformed low education ones only in WR, $t(147) = -3$; $p = .004$; $d = .51$. On the contrary, in the middle age group high education participants outperformed low education ones in both WR, $t(134) = -2.8$; $p = .006$; $d = .49$ and CW, $t(134) = -2.6$; $p = .009$; $d = .48$, while differences between groups in CN were only marginally significant, $t(134) = -1.9$; $p = .06$; $d = .33$. Accordingly, only the middle-age group was divided into two education levels for stratification of norms (0–11 and 12+ years).

The SCH sample was divided into two groups of education (see Table 5 for descriptive statistics): low and high level of education. Both groups differed in WR score, $t(147) = -3.80$; $p < .001$; $d = .64$; in CN score, $t(147) = -2.89$; $p < .004$; $d = .49$; in CW score, $t(147) = -3.92$; $p < .001$; $d = .71$; and IS score, $t(147) = -2.28$; $p < .024$; $d = .41$.

The TBI sample was divided into two groups of education (see Table 6 for descriptive statistics): low and high level of education. Both groups differed in WR, $t(156) = -4.1$; $p < .001$; $d = .67$; in CN, $t(156) = -2.91$; $p = .004$; $d = .47$; and in CW scores, $t(156) = -3.2$; $p = .002$; $d = .53$ but not in IS score, $t(156) = -1.1$; $p = .277$; $d = .19$.

Tables 4–6 provide descriptive statistics (mean, standard deviation, maximum, minimum, skewness, and kurtosis) for age, education, and Stroop variables in the three different samples. Tables 7–9 provide normative data for Stroop variables stratified by age and education in the case of HC, and by education in the case of SCH, and TBI samples.

Table 4. Statistical properties of the demographic and Stroop variables for healthy controls ($n = 285$)

Young group $N = 149$; male = 66; female = 83						
	<i>M</i>	<i>SD</i>	Min	Max	Skewness	Kurtosis
Age	23.8	4	15	31	–	–
Edu	13.1	3.3	6	23	–	–
WR	108.8	15.2	53	147	–.5	.7
CN	81.1	13.6	41	113	–.3	–.3
CW	55.2	11.6	17	84	–.2	.6
IS	9	8.9	–25.7	28.8	–.4	1.1
Middle-age group; Low Education 0–11 years $N = 69$; male = 17; female = 52						
	<i>M</i>	<i>SD</i>	Min	Max	Skewness	Kurtosis
Age	58.2	10.8	33	78	–	–
Edu	8.6	1.5	2	11	–	–
WR	95.6	17	62	140	–.4	.2
CN	66.2	12	41	95	.6	.1
CW	37.6	9.7	22	70	1	1.4
IS	–1.2	8.5	–16.8	21.3	.6	.2
Middle-age group; High Education 12+ years $N = 67$; male = 32; female = 35						
	<i>M</i>	<i>SD</i>	Min	Max	Skewness	Kurtosis
Age	53.3	13.5	32	80	–	–
Edu	15.6	2.6	12	24	–	–
WR	103.7	16.6	54	143	–.6	.6
CN	70.5	14.6	20	104	–.6	1.3
CW	42.5	11.6	16	68	0	–.5
IS	0.7	9	–20.3	22.3	–.1	–.4

Note: Edu = Education in years; WR = Word Reading; CN = Color Naming; CW = Color Word; IS = Interference Score.

Discussion

The increase in the number of studies providing Stroop norms from different non-clinical populations has helped to avoid the risk of under- or over-estimating cognitive functioning due to cultural or socio-demographical differences (Norman et al., 2011; Seo et al., 2008; Van der Elst et al., 2006; Zalonis et al., 2009). Clinical norms have been also provided for different psychological tools such as Wisconsin Card Sorting Tests (Iverson et al., 2000), Trail Making Test (Periáñez et al., 2007), Rey's Auditory Verbal Learning Test (Badcock et al., 2011), or Boston Naming Test (Casals-Coll et al., 2014). The aim of this study was to provide clinical norms of the Golden's (1999) version of the Stroop test for Spanish patients with TBI and

Table 5. Statistical properties of demographic and Stroop variables for patients with schizophrenia ($n = 149$)

Low Education 0–10 years $N = 88$; male = 57; female = 31						
	Mean	<i>SD</i>	Min	Max	Skewness	Kurtosis
Age	30.3	10.9	16	58	–	–
Edu	7.8	1.4	6	10	–	–
WR	88.7	16.5	45	125	–.1	–.1
CN	59	10.6	31	85	–.2	.2
CW	31.8	9	7	55	–.2	.4
IS	–3.4	6.4	–22.6	13.6	–.3	.7
High Education 11+ years $N = 61$; male = 25; female = 36						
	Mean	<i>SD</i>	Min	Max	Skewness	Kurtosis
Age	34	9.9	17	60	–	–
Edu	14	2.2	12	17	–	–
WR	99.3	17.2	59	140	0	–.4
CN	64.9	14.3	28	117	.4	2.3
CW	38.3	11	17	76	.8	1.5
IS	–.7	7.8	–21.5	17.1	–.2	–.1

Note: Edu = Education in years; WR = Word Reading; CN = Color Naming; CW = Color Word; IS = Interference Score.

schizophrenia, together with matching data for healthy individuals. In the following sections differences between groups and criteria for stratification will be discussed in relation to preceding literature.

Results from the ANCOVAs comparing Stroop WR and CN scores between the three samples showed that both groups of patients scored lower than healthy controls in all test conditions. Moreover, both TBI and SCH groups exhibited greater interference effects than healthy controls, as reflected by the IS. Regarding TBI patients, the results confirmed those from prior investigations suggesting that TBI is associated with a generalized slowing in task performance across all test conditions (Felmingham et al., 2004; Ríos et al., 2004). In addition, differences in IS between TBI patients and healthy controls indicated that the TBI group displayed difficulties in the interference condition of the test. In fact, 42 % of the TBI participants obtained a negative score. These results support preceding studies suggesting that, in addition to slowed information processing speed, TBI might be associated to a deficit in selective attention (Summers, 2006). Regarding SCH patients, our data agree with previous findings showing that slowness is a prevalent feature in this population when facing the Stroop task (Brébion et al., 2000; Rodríguez-Sánchez et al., 2007). However, attention deficits also seem to play a role that would account

Table 6. Statistical properties of demographic and Stroop variables for patients with (n = 158)

Low Education 0–12 years N = 82; male = 72; female = 10						
	Mean	SD	Min	Max	Skewness	Kurtosis
Age	28.8	11.7	15	62	–	–
Edu	9.4	1.6	7	12	–	–
WR	72.5	23.2	8	122	–.4	–.3
CN	52.4	16.1	11	85	–.3	–.4
CW	31	12.9	1	62	.2	–.1
IS	.9	8.8	–27	19.6	–.2	.4

High Education 13+ years N = 76; male = 55; female = 21						
	Mean	SD	Min	Max	Skewness	Kurtosis
Age	35.9	12.2	18	72	–	–
Edu	16.1	1.9	13	20	–	–
WR	86.5	20.2	37	130	–.4	.1
CN	59.9	16.2	19	100	–.2	.1
CW	37.6	13.3	10	85	.4	1
IS	2.5	10.4	–30.5	54	1	8

Note: Edu = Education in years; WR = Word Reading; CN = Color Naming; CW = Color Word; IS = Interference Score.

for the presence of differences between SCH patients and healthy controls in IS (Westerhausen, Kompus, & Hugdahl, 2011). Moreover, the comparison between both clinical samples revealed that, even when TBI patients were significantly slower than SCH patients (as revealed by differences in WR and CN conditions), SCH patients scored lower than TBI in IS. Thus, although results revealed the presence of a specific deficit in executive control in both clinical samples, SCH patients displayed the greatest interference effect as compared to TBI (66% of the SCH patients obtained a negative IS).

The analyses derived from the stratification of HC showed that age was the best predictor of individual's performance in the Stroop test. In fact, it accounted for a significant portion of variance in all test conditions. Different authors have recognized that ageing involve a slowing in CN as well as an increase in the interference effect (Moering et al., 2004; Van der Elst et al., 2006). Contrasting these results and those from the present study, Rognoni et al. (2013) found that age did not have any effect on Stroop scores in their Spanish sample. However, it has to be noted that their sample of healthy Spanish controls did not include participants over 50 years old. This fact may account for this apparent inconsistency regarding the role of age, since Stroop effects seem to be more evident in the last

Table 7. Percentile ranks for healthy controls (n = 285) stratified by age and education

Young group 15–31 years (n = 149)				
	WR	CN	CW	IS
Perc				
5	80	60	36	–6.4
10	89	64	42	–3.3
15	94	67	45	.9
20	98	70	46	2.4
25	100	71	48	4.7
30	101	73	48	5.3
35	103	75	50	6.4
40	105	77	52	7.1
45	107	79.5	54	8.4
50	109	81	55	9.2
55	112	83	57	9.9
60	114	86	58	11
65	117	88.5	60	12.1
70	118	91	61	12.8
75	120	92	63	14.7
80	122	93	64	16.3
85	124	96	67	18.1
90	127	98	69	18.8
95	133	101	75	25.1

Middle age group 32–80 years (n = 136)

	WR		CN		CW		IS	
Edu	0–11	12+	0–11	12+	0–11	12+	0–11	12+
Perc	n = 69	n = 67	n = 69	n = 67	n = 69	n = 67	n = 69	n = 67
5	69	70.8	47.5	46.8	23.5	23.8	–13.5	–15.2
10	73	80	55	53.6	26	26.6	–11.8	–12.1
15	75.5	86	56	54.4	28	30	–8.6	–9.6
20	80	87.6	56	57.6	29	32	–7.6	–6.1
25	83	94	57	61	30	33	–7.3	–5.1
30	84	97	60	65	32	34.4	–6.3	–3.5
35	89.5	100	60.5	67	33	37.6	–5.4	–2.6
40	92	102.2	61	69	35	40.2	–5	–2.2
45	94	104.6	63	70	36	42	–4.1	–.4
50	97	107	64	72	37	43	–3.4	.2
55	99.5	109.4	66.5	74	38	44	–.7	2
60	100	110	68	75	39	46	.9	3.3
65	103	111.2	69	76	40	48	1.7	4.6
70	104	112	70	78.6	41	48.6	2.6	6.1
75	105	114	71.5	80	43	49	3.8	7.1
80	107	116.4	75	81.4	45	53.4	5.1	8.6
85	110	119.6	80	84.6	46.5	55	6.9	11.1
90	115	122.2	87	88	50	57.2	10.8	12.6
95	131	128.4	92	94	56.5	63.6	16.5	15.3

Note: Edu = Education in years; Perc = Percentile; WR = Word Reading; CN = Color Naming; CW = Color Word; IS = Interference Score.

Table 8. Percentile ranks for the schizophrenia sample (n = 149) stratified by education

	WR		CN		CW		IS	
	0-10	11+	0-10	11+	0-10	11+	0-10	11+
Edu	0-10	11+	0-10	11+	0-10	11+	0-10	11+
Perc	n = 88	n = 61	n = 88	n = 61	n = 88	n = 61	n = 88	n = 61
5	60.5	71.3	37.8	39.2	16.4	20.1	-15.8	-13.7
10	69	78	45	46.2	21	24.2	-11.3	-10.8
15	73.4	80	48	51.9	23	28	-10.2	-9.5
20	75	81	50	56	24	29.4	-8	-8
25	78	87	51.3	57	26	30	-6.7	-6.3
30	80	90.6	53	60	28	33.6	-5.6	-4.6
35	82.2	92.7	56	60	28	34.7	-4.9	-3.2
40	83.6	94.8	57	61.8	29	35	-4.5	-1.6
45	85.1	96	59.1	62.9	31	36.9	-3.7	-1.3
50	86.5	100	60	65	32	38	-3	-8
55	89	102	61	66	33	38.1	-2.6	-3
60	93	103.2	62	68	35	40	-1.7	.9
65	94.9	105.3	63	69	35.9	40	-1	2.8
70	100	109.6	65	70	37	41.4	-.6	3.1
75	100.8	113	66	73	38	44	.5	4.7
80	104	114.6	67	75	39.2	45.6	2.6	6.3
85	106	118.1	69.7	77.7	40	48.4	3	8.6
90	110	121.6	71.1	82.8	43.1	54	4.3	9.9
95	118.2	130.4	75	87.7	46	59.7	6.5	12.1

Note: Edu = Education in years; Perc = Percentile; WR = Word Reading; CN = Color Naming; CW = Color Word; IS = Interference Score.

decades of life (Uttl & Graf, 1997). Like age, education resulted to be a good predictor of HC Stroop performance in the present study, accounting for significant portions of variance in both WR and CN scores. The demographic effects of education have been consistently reported for both Spanish and non-Spanish populations (Moering et al., 2004; Peña-Casanova et al., 2009; Rognoni et al., 2013; Van der Elst et al., 2006). In the present work, sex did not have an influence in any Stroop score. Although in some works women have tended to score higher in color-naming (Moering et al., 2004; Van der Elst et al., 2006), sex differences on the CW condition are not always present (Golden & Freshwater, 2002; Moering et al., 2004; Rognoni et al., 2013). Taken together, these results highlight the importance of considering norms reflecting the specific impact of demographic variables in different populations.

Similarly to HC, both age and education impacted Stroop scores in the two clinical groups. The analyses revealed that education was a good demographic predictor of Stroop performance in both TBI and SCH samples. Age accounted for a portion of the variance of CW and IS in the SCH sample, and of CN, CW, and IS in the TBI sample. However, this portion was in most

Table 9. Percentile ranks for the TBI sample (n = 158) stratified by education

	WR		CN		CW		IS	
	0-12	13+	0-12	13+	0-12	13+	0-12	13+
Edu	0-12	13+	0-12	13+	0-12	13+	0-12	13+
Perc	n = 82	n = 76	n = 82	n = 76	n = 82	n = 76	n = 82	n = 76
5	30.1	48.6	24	32.7	11	14.4	-14.5	-12.4
10	40	54.4	29	37.7	13.3	23.7	-9	-9.5
15	48.4	65	33.4	41.2	15.5	25	-7.3	-7.4
20	51	71.2	40	45	18.6	26	-5.8	-4.1
25	55	75	42	49.3	21.5	27.3	-5	-3.1
30	56	79	45	52	24	29	-3.2	-1.5
35	62.1	82	46	55	27	31	-3	.1
40	69	83.8	49.2	58	29	33	-2.6	.8
45	73.4	86	51.4	59.7	30	34	-1.4	1.7
50	78.5	88	54	61	31	38	.2	2.9
55	80.7	91	55	63.4	32.7	39.4	1.7	4.2
60	81	92.2	57	64.4	34.8	42	2.3	5
65	83	96	60	67	36	45	4	5.6
70	88.2	97	64	68	37	45	6.3	6.9
75	92	100	65.3	71.8	38.3	47	7.4	7.9
80	93	100	67	73	40.4	49	8.8	9.5
85	95	107.5	70	75	42.6	50.5	10.5	10.4
90	99.4	111.3	71	80	50.5	54	13.5	12.2
95	105.7	121.3	77.7	85.2	55.9	58.2	14.7	15.3

Note: Edu = Education in years; Perc = Percentile; WR = Word Reading; CN = Color Naming; CW = Color Word; IS = Interference Score.

cases inferior to that accounted by education. This result fits prior TBI and SCH clinical norms where education was the main variable selected for stratification in a different attentional test (Periáñez et al., 2007). Thus, it is important to consider these demographic variables when interpreting Stroop performance in the clinical context. Ignoring their influence could lead to misinterpretations when monitoring evolution or recovery, over- or underestimating patients' difficulties.

Regarding the generalization of the present data, it seems plausible that the influence of TBI and SCH is higher than the influence of cultural and demographic variables. The effect sizes found when comparing HC, TBI and SCH groups are higher than those found when studying differences due to age and education. It is also known that ethnicity (Moering et al. 2004; Norman et al., 2011), country (Buré-Reyes et al., 2013), and language (Rosselli et al., 2002) may impact test scores (Strauss et al., 2006). However, no cultural or demographic variables show higher effect sizes than those found among healthy and pathological groups.

To summarize, the major value of the present study was to provide a set of clinical norms to determine more precisely the extent to which Stroop scores on WR, CN, CW, and IS reflect impairments in

performance, and changes across time. This issue has implications for research, forensic, and clinical settings allowing a more precise description of patients, and a more sensible detection of changes in performance across time. However, it should be clearly established that the present clinical norms do not avoid the risk of misinterpreting undesired retest effects (i.e., practice effects or procedural learning) as real cognitive changes. Complementary normative methods have been recently proposed to solve this issue, and should be applied before using the present clinical norms in longitudinal assessments (Calamia, Markon, & Tranel, 2013; Van del Elst et al., 2013).

References

- Arfanakis K., Haughton V. M., Carew J. D., Rogers B. P., Dempsey R. J., & Meyerand M. E. (2002). Diffusion tensor MR imaging in diffuse axonal injury. *American Journal of Neuroradiology*, 23, 794–802.
- Badcock J. C., Dragovic M., Dawson L., & Jones R. (2011). Normative data for Rey's Auditory Verbal Learning Test in individuals with schizophrenia. *Archives of Clinical Neuropsychology*, 26, 205–213. <http://dx.doi.org/10.1093/arclin/acr005>
- Bur -Reyes A., Hidalgo-Ruzzante N., Vilar-L pez R., Gontier J., S nchez L., P rez-Garc a M., & Puente A. E. (2013). Neuropsychological test performance of Spanish speakers: Is performance different across different Spanish-speaking subgroups? *Journal of Clinical and Experimental Neuropsychology*, 35, 404–412. <http://dx.doi.org/10.1080/13803395.2013.778232>
- Br bion G., Smith M. J., Gorman J. M., Malaspina D., Sharif Z., & Amador X. (2000). Memory and schizophrenia: Differential link of processing speed and selective attention with two levels of encoding. *Journal of Psychiatric Research*, 34, 121–127. [http://dx.doi.org/10.1016/S0022-3956\(99\)00050-3](http://dx.doi.org/10.1016/S0022-3956(99)00050-3)
- Calamia M., Markon K., & Tranel D. (2013). Scoring higher the second time around: Meta-analyses of practice effects in neuropsychological assessment. *The Clinical Neuropsychologist*, 27, 1077–1105. <http://dx.doi.org/10.1080/13854046.2013.809795>
- Casals-Coll M., S nchez-Benavides G., Meza-Cavazos S., Manero R. M., Aguilar M., Badenes D., ... NEURONORMA Study Team. (2014). Spanish multicenter normative studies (NEURONORMA project): Normative data and equivalence of four BNT short-form versions. *Archives of Clinical Neuropsychology*, 29, 60–74. <http://dx.doi.org/10.1093/arclin/act085>
- Coubard O. A., Ferruffino L., Boura M., Gripon A., Renaud M., & Bherer L. (2011). Attentional control in normal aging and Alzheimer's disease. *Neuropsychology*, 25, 353–367. <http://dx.doi.org/10.1037/a0022058>
- Crespo-Facorro B., P rez-Iglesias R., Ramirez-Bonilla M., Mart nez-Garc a O., Llorca J., & V zquez-Barquero J. L. (2006). A practical clinical trial comparing haloperidol, risperidone, and olanzapine for the acute treatment of first-episode nonaffective psychosis. *The Journal of Clinical Psychiatry*, 67, 1511–1521. <http://dx.doi.org/10.4088/jcp.v67n1004>
- Faul F., Erdfelder E., Lang A. G., & Buchner A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. <http://dx.doi.org/10.3758/BF03193146>
- Felmingham K. L., Baguley I. J., & Green A. M. (2004). Effects of diffuse axonal injury on speed of information processing following severe traumatic brain injury. *Neuropsychology*, 18, 564–571. <http://dx.doi.org/10.1037/0894-4105.18.3.564>
- Golden C. (1978). *Stroop Color and Word Test: Manual for clinical and experimental uses*. Chicago, IL: Stoelting.
- Golden C. J. (1999). *Stroop, Test de Colores y Palabras. Manual de aplicaci n*. [Stroop Color and Word Test. Examiner's manual]. Madrid, Spain: TEA Ediciones.
- Golden C. J., & Freshwater S. M. (2002). *The Stroop Color and Word test: A manual for clinical and experimental uses*. Chicago, IL: Stoelting Co.
- Gonz lez-Blanch C.,  lvarez-Jim nez M., Rodr guez-S nchez J. M., P rez-Iglesias R., V zquez-Barquero J. L., & Crespo-Facorro B. (2006). Cognitive functioning in the early course of first-episode schizophrenia spectrum disorders: Timing and patterns. *European Archives of Psychiatry and Clinical Neuroscience*, 256, 364–371. <http://dx.doi.org/10.1007/s00406-006-0646-6>
- Hepp H. H., Maier S., Hermle L., & Spitzer M. (1996). The Stroop effect in schizophrenic patients. *Schizophrenia Research*, 22, 187–195. [http://dx.doi.org/10.1016/S0920-9964\(96\)00080-1](http://dx.doi.org/10.1016/S0920-9964(96)00080-1)
- Iverson G. L., Slick D. J., & Franzen M. D. (2000). Clinical normative data for the WCST-64 following uncomplicated mild head injury. *Applied Neuropsychology*, 7, 247–251. http://dx.doi.org/10.1207/s15324826an0704_7
- Jensen A. R., & Rohwer Jr W. D. (1966). The Stroop Color-Word Test: A review. *Acta Psychologica*, 25, 36–93. [http://dx.doi.org/10.1016/0001-6918\(66\)90004-7](http://dx.doi.org/10.1016/0001-6918(66)90004-7)
- Kizilbash A., Warschusky S., & Donders J. (2001). Assessment of speed of processing after pediatric head trauma: Need for better norms. *Pediatric Rehabilitation*, 4, 71–74. <http://dx.doi.org/10.1080/13638490026403>
- Lansbergen M. M., Kenemans J. L., & van Engeland H. (2007). Stroop interference and attention-deficit/hyperactivity disorder: A review and meta-analysis. *Neuropsychology*, 21, 251–262. <http://dx.doi.org/10.1037/0894-4105.21.2.251>
- Levin H. S., O'Donnell V. M., & Grossman R. G. (1979). The Galveston Orientation and Amnesia Test. A practical scale to assess cognition after head injury. *Journal of Nervous and Mental Disease*, 167, 675–684. <http://dx.doi.org/10.1037/t28509-000>
- Mayas J., Fuentes L. J., & Ballesteros S. (2012). Stroop interference and negative priming (NP) suppression in normal aging. *Archives of Gerontology and Geriatrics*, 54, 333–338. <http://dx.doi.org/10.1016/j.archger.2010.12.012>
- Moering R. G., Schinka J. A., Mortimer J. A., & Graves A. B. (2004). Normative data for elderly African Americans for

- the Stroop Color and Word Test. *Archives of Clinical Neuropsychology*, 19, 61–71. [http://dx.doi.org/10.1016/S0887-6177\(02\)00219-6](http://dx.doi.org/10.1016/S0887-6177(02)00219-6)
- Norman M. A., Moore D. J., Taylor M., Franklin D. Cysique L., Ake C., ... The HNRC Group** (2011). Demographically corrected norms for African Americans and Caucasians on the Hopkins Verbal Learning Test-Revised, Brief Visuospatial Memory Test-Revised, Stroop Color and Word Test, and Wisconsin Card Sorting Test 64-Card Version. *Journal of Clinical and Experimental Neuropsychology*, 33, 793–804. <http://dx.doi.org/10.1080/13803395.2011.559157>
- Peña-Casanova J., Quiñones-Ubeda S., Gramunt-Fombuena N., Quintana M., Aguilar M., Molinuevo J. L., ... NEURONORMA Study Team** (2009). Spanish Multicenter Normative Studies (NEURONORMA Project): Norms for the Stroop color-word interference test and the Tower of London-Drexel. *Archives of Clinical Neuropsychology*, 24, 413–429. <http://dx.doi.org/10.1093/arclin/acp043>
- Periáñez J. A., Ríos-Lago M., Rodríguez-Sánchez J. M., Adrover-Roig D., Sánchez-Cubillo I., Crespo-Facorro B., ... Barcelo F.** (2007). Trail Making Test in traumatic brain injury, schizophrenia, and normal ageing: Sample comparisons and normative data. *Archives of Clinical Neuropsychology*, 22, 433–447. <http://dx.doi.org/10.1016/j.acn.2007.01.022>
- Ríos M., Periáñez J. A., & Muñoz-Céspedes J. M.** (2004). Attentional control and slowness of information processing after severe traumatic brain injury. *Brain Injury*, 18, 257–272. <http://dx.doi.org/10.1080/02699050310001617442>
- Rodríguez-Sánchez J. M., Crespo-Facorro B., González-Blanch C., Perez-Iglesias R., Vázquez-Barquero J. L., & PAFIP Group Study.** (2007). Cognitive dysfunction in first episode psychosis: The processing speed hypothesis. *The British Journal of Psychiatry*, 191, 107–110. <http://dx.doi.org/10.1192/bjp.191.51.s107>
- Rognoni T., Casals-Coll M., Sánchez-Benavides G., Quintana M., Manero R. M., Calvo L., ... Peña-Casanova J.** (2013). Estudios normativos españoles en población adulta joven (proyecto NEURONORMA jóvenes): Normas para las pruebas Stroop Color-Word Interference Test y Tower of London-Drexel University. [Spanish normative studies in young adults (NEURONORMA young adults project): norms for Stroop Color-Word Interference and Tower of London-Drexel University Tests]. *Neurología*, 28, 73–80. <http://dx.doi.org/10.1192/bjp.191.51.s107>
- Rosselli M., Ardila A., Santisi M. N., Arecco A. D. R., Salvatierra J., Conde A., & Lenis B.** (2002). Stroop effect in Spanish-English bilinguals. *Journal of the International Neuropsychological Society*, 15, 692–693. [http://dx.doi.org/10.1016/s0887-6177\(00\)80077-3](http://dx.doi.org/10.1016/s0887-6177(00)80077-3)
- Seo E. H., Lee D. Y., Choo I. H., Kim S. G., Kim K. W., Youn J. C., ... Woo J. I.** (2008). Normative study of the Stroop Color and Word Test in an educationally diverse elderly population. *International Journal of Geriatric Psychiatry*, 23, 1020–1027. <http://dx.doi.org/10.1002/gps.2027>
- Strauss E., Sherman E. M. S., & Spreen O.** (2006). *A compendium of Neuropsychological Tests: Administration, norms, and commentary* (Third Edition Ed.). New York, NY: Oxford University Press.
- Summers M. J.** (2006). Increased inattentive blindness in severe traumatic brain injury: Evidence for reduced distractibility? *Brain Injury*, 20, 51–60. <http://dx.doi.org/10.1080/02699050500309742>
- Teasdale G., & Jennett B.** (1974). Assessment of coma and impaired consciousness. A practical scale. *Lancet*, 304, 81–84. [http://dx.doi.org/10.1016/S0140-6736\(74\)91639-0](http://dx.doi.org/10.1016/S0140-6736(74)91639-0)
- Uttl B., & Graf P.** (1997). Color-Word Stroop test performance across the adult life span. *Journal of Clinical and Experimental Neuropsychology*, 19, 405–420. <http://dx.doi.org/10.1080/01688639708403869>
- Van der Elst W., Van Boxtel M. P. J., Van Breukelen G. J. P., & Jolles J.** (2006). The Stroop Color-Word Test: Influence of age, sex, and education; and normative data for a large sample across the adult age range. *Assessment*, 13, 62–79. <http://dx.doi.org/10.1177/1073191105283427>
- Van der Elst W., Molenberghs G., Van Boxtel M. P. J., & Jolles J.** (2013). Establishing normative data for repetitive cognitive assessment: A comparison of different statistical methods. *Behavior Research Methods*, 45, 1073–1086. <http://dx.doi.org/10.3758/s13428-012-0305-y>
- Westerhausen R., Kompus K., & Hugdahl K.** (2011). Impaired cognitive inhibition in schizophrenia: A meta-analysis of the Stroop interference effect. *Schizophrenia Research*, 133, 172–181. <http://dx.doi.org/10.1016/j.schres.2011.08.025>
- Zalonis I., Christidi F., Bonakis A., Kararizou E., Triantafyllou N. I., Paraskevas G., ... Vasilopoulos D.** (2009). The stroop effect in Greek healthy population: Normative data for the Stroop Neuropsychological Screening Test. *Archives of Clinical Neuropsychology*, 24, 81–88. <http://dx.doi.org/10.1093/arclin/acp011>