This is an accepted manuscript of an article published by Taylor & Francis in Developmental Neuropsychology on 23 Dec 2020 available online: https://doi.org/10.1080/87565641.2020.1863407

Multivariate Base Rates of Low Scores on Tests of Executive Functions in a Multi-Country Latin American Sample

Diego Rivera
Departamento de Ciencias de la Salud, Universidad Pública de Navarra, Pamplona, Spain
diegoriveraps@gmail.com

Guido Mascialino
Escuela de Psicología, Universidad de Las Américas, Quito, Ecuador
guido.mascialino@udla.edu.ec

Brian L. Brooks
Neuropsychology Service, Alberta Children's Hospital, Calgary, Alberta, Canada;
Departments of Pediatrics, Clinical Neurosciences, and Psychology, University of Calgary, Calgary, Alberta, Canada;
Alberta Children's Hospital Research Institute, University of Calgary, Calgary, Alberta, Canada
Brian.Brooks@albertahealthservices.ca

Laiene Olabarrieta-Landa
Departamento de Ciencias de la Salud, Universidad Pública de Navarra, Pamplona, Spain
laieneolabarrieta@gmail.com

Melina Longoni
Clínica De Rehabilitación Rêverie, Buenos Aires, Argentina.
melinalongoni@hotmail.com

Javier Galarza-del-Angel
Facultad de Ciencias Humanas, Universidad Autónoma de Baja California, Mexicali, Mexico
javiergalarza@uabc.edu.mx

Juan Carlos Arango-Lasprilla
Biocruces Bizkaia Health Research Institute. Cruces University Hospital. Barakaldo, Spain
IKERBASQUE. Basque Foundation for Science. Bilbao, Spain
jcalasprilla@gmail.com

Address to correspondence:
Juan Carlos Arango Lasprilla, Ph.D
Biocruces Bizkaia Health Research Institute. Cruces University Hospital
IKERBASQUE. Basque Foundation for Science.
Plaza de Cruces s/n. 48903
Barakaldo. Bizkaia, Spain.
Phone: (34) 946006000 (Ext. 7963)
E mail: jcalasprilla@gmail.com

Abstract

Objective: Knowing the prevalence of low scores on neuropsychological tests that are administered jointly can help avoid erroneous interpretations of test scores. The objective of the study was to determine the prevalence of low scores in a Latino population for two neuropsychological commonly used to evaluate executive functions and to compare the number of low scores obtained using normative data from a Spanish-speaking population versus an English-speaking population. **Method**: Healthy adults (N=5402) from 12 countries in Latin-America were recruited and administered the Modified Wisconsin Card Sorting Test (M-WCST) and Stroop Color-Word Interference Test. Two-thirds were women, and the average age was 53.5±20.0 (range 18-95) years. Each participant was categorized based on his/her number of low scoring tests in specific percentile cutoff groups 25^{th}, 16^{th}, 10^{th}, 5^{th}, and 2^{nd}. **Results**: Between 32.8% (Puerto Rico) and 48.6% (Guatemala) scored below the 16^{th} percentile on one or more scores. Between 22.7% (Paraguay) and 34.8% (El Salvador) scored below the 10^{th} percentile on at least 1 of the 5 scores. Between 12.9% (Puerto Rico) and 20.3% (Honduras) scored below the 5^{th} percentile on one or more scores. In addition, the number of low scores at each percentile cutoff was higher when estimated using normative data from an English-speaking population versus a Spanish-speaking population. **Conclusions**: Consistent with existing research on interpretation of multiple tests administration, having low scores on measures of executive functioning can be common. Clinicians working with Spanish-speaking adults should take into account the higher probability of low scores on measures of executive functions to reduce false-positive diagnoses of cognitive deficits in an individual.

Introduction

Executive functions are those cognitive abilities that allow a person to engage in self-directed purposeful behavior (Baddeley, 1996; Lezak, 2012a; Shallice et al., 1996). There are differing views on the precise nature of executive functions, but most authors describe them as composed of several interrelated subprocesses (Jurado & Rosselli, 2007). Core executive skills include inhibitory control, working memory, and cognitive flexibility (Diamond, 2013), but some authors also include behavioral dimensions such as volition, purposive action, and decision-making as essential aspects as well (Lezak, 2012a). While there is some disagreement on the unitary versus heterogenous nature of these abilities, evidence from neuropsychological and imaging studies suggest there is both unity and diversity in their functioning (Friedman & Miyake, 2017). Executive functions typically are associated with frontal lobe functioning and executive networks, and include the capacity for establishing goals, developing plans, and executing them in an effective manner (Hofmann et al., 2012; Stuss, 2011). As such, they are essential for adaptive behavior and independent functioning, which makes them necessary domains to assess in a neuropsychological evaluation.

Proper measurement of executive functioning, including the appropriate interpretation of test results using advanced psychometric methods, is essential for many diagnostic considerations across the lifespan, including (but not limited to) developmental disorders (e.g., autism spectrum disorder, fetal alcohol spectrum disorder, attention disorders, epilepsy, Tourette's syndrome), acquire brain injuries (e.g., stroke, traumatic brain injury, toxic exposure), psychiatric disorders (e.g., schizophrenia, depression, post-traumatic stress disorder, anxiety disorders, bipolar disorder), brain tumors, and dementias (e.g., frontal-temporal dementia, Lewy Body dementia, vascular dementia) (Baron, 2018; Horton & Wedding, 2008; Kolb &

Whishaw, 2015; Morgan & Ricker, 2018; Yeates et al., 2010). There are a wide variety of neuropsychological instruments to assess executive functions (Lezak, 2012b; Strauss et al., 2006; Rabin et al., 2016), among them are the Modified Wisconsin Card Sorting Test (M-WCST; Nelson, 1976; Schretlen, 2010) and the Stroop Color and Word Test (Golden & Freshwater, 2002; Stroop, 1935), both of which are commonly used.

The M-WCST is a test that measures abstract thought, cognitive set shifting, learning from feedback, goal-oriented behavior, and response inhibition (Nelson, 1976; Strauss et al., 2006). The M-WCST has been used to investigate multiple pathologies in studies, including depression (Fossati et al., 2001), schizophrenia (Hartman et al., 2003), mild cognitive impairment (Borkowska et al., 2009), Huntington's disease (Peinemann et al., 2005), Alzheimer's disease (Bondi et al., 1993), and Parkinson's disease (Gotham et al., 1988). It has also been utilized in several countries including Argentina, Bolivia, Chile, Cuba, El Salvador, Guatemala, Honduras, Mexico, Paraguay, Peru, Puerto Rico, Brazil, Spain, Italy, and Japan (Arango-Lasprilla et al., 2015; Cianchetti et al., 2007; del Pino et al., 2016; Kado et al., 2004; Nakayama et al., 1990; Zimmermann et al. , 2015).

The Stroop test (Stroop, 1935) is a measure of executive functions that assesses cognitive control, selective attention, and response inhibition (Lezak, 2012a; Strauss et al., 2006). The test requires naming colors, reading color words (e.g. red, green), and lastly naming the color of the ink in which a color word is printed when these are mismatched. There have been various versions of this test, which vary by number of trials, colors utilized, and type of stimuli (Strauss et al., 2006).

The Stroop Color and Word Test has also been utilized globally in countries such as Argentina, Bolivia, Chile, Cuba, El Salvador, Guatemala, Honduras, Mexico, Paraguay, Peru, Puerto Rico, Italy, Spain, the Netherlands, Portugal, South Africa, Japan, and India (Brugnolo et al., 2016; Ikeda et al., 2011; Llinas-Regla et al., 2013; Maqbool et al., 2015; Paulo et al., 2011; Rivera et al., 2015; Savitz & Jansen, 2003).

Normative data on these tests were obtained predominantly in English-speaking regions; however, recent studies have obtained normative data for Spanish-speaking adults for both the M-WCST and the Stroop Color and Word Test (Arango-Lasprilla et al., 2015; del Pino et al., 2016; Peña-Casanova et al., 2009; Rivera et al., 2015; Rognoni et al., 2013). The availability and utilization of proper norms reduces the potential for misinterpretation of scores when evaluating this population.

Another common source of score misinterpretation in neuropsychological assessment is to draw inferences about cognitive performance based on single scores of a neuropsychological battery rather than performance across all measures simultaneously (Binder et al., 2009; Brooks & Iverson, 2010; Ingraham, 1996; Schretlen et al., 2008). Statistically, the probability of obtaining an impaired score on one measure increases as the number of tests administered also increases, a fact proven empirically through multiple studies investigating the base rates of low scores, also called the multivariate base rates (MVBRs), when administering multiple tests to healthy individuals (Binder et al., 2009; Brooks et al., 2013; Brooks & Iverson, 2010; Brooks et al., 2012; Brooks et al., 2013; Cook et al., 2019; Holdnack et al., 2017; Karr et al., 2017). Most of the data on base rates of low scores in a healthy population has been collected in English-speaking countries and, to the best of the authors' knowledge, there is data on MVBRs of low scores on Spanish-speakers in language (Olabarrieta-Landa et al., 2019) and memory (Rivera et al., 2019) domains, however there is no data on MVBRs of low scores focused in executive function on Spanish-speakers.

This study will examine MVBRs in a Spanish-speaking adult Latino sample across 11 countries and the commonwealth of Puerto Rico who completed the Stroop

Color and Word Test and the M-WCST and to compare the number of low scores obtained using normative data from a Spanish-speaking population versus an English-speaking population. Filling this vacuum in the literature will further limit the potential for misinterpretation of findings in the neuropsychological evaluation of this population. In order to facilitate this, an appendix with low scores MVBRs for simultaneous administration of the tests will be included.

Method

Participants

The sample consisted of 5,402 healthy individuals who were recruited from Argentina, Bolivia, Chile, Colombia, Cuba, El Salvador, Guatemala, Honduras, Mexico, Paraguay, Peru, and Puerto Rico. The demographic characteristics (age, education, and sex) by country can be found in Table 1.

Insert Table 1

To be eligible for study participation, individuals must have met the following requirements: (a) were between 18 to 95 years of age, (b) were born and currently live in the country where the protocol was conducted, (c) spoke Spanish as their native language, (d) had completed at least one year of formal education, (e) were able to read and write at the time of evaluation, (f) scored ≥ 23 on a Spanish version of the Mini-Mental State Examination (MMSE; Folstein et al., 1975; Villaseñor-Cabrera et al., 2010), (g) scored ≤ 4 on a Spanish version of the Patient Health Questionnaire–9 (PHQ-9; Kroenke et al., 2001), and (h) scored ≥ 90 on the Barthel Index (Mahoney & Barthel, 1965).

A self-report questionnaire was administered to collect information about the participants' medical history and health status. Participants were excluded if they reported or endorsed the following: (a) medical services received for diagnosed neurological or psychiatric conditions, (b) daily consumption and/or use of an illicit substance, (c) history of chronic disease (e.g., diabetes mellitus), (d) regular use of pain or other medications that may impact cognitive functioning, and/or (e) severe visual and/or hearing deficit. All participants were community volunteers who did not receive financial compensation for participation.

Measures

The Modified Wisconsin Card Sorting Test (M-WCST). It consists of four stimulus cards and 48 response cards. Each card varies in shape (cross, circle, triangle, or star), color (red, blue, yellow, or green), and number (one to four). The objective is to classify correctly the stimulus card according to certain rule until completion of a category. The test continues until all six categories are classified or until the whole volume has been used (Schretlen, 2010). The test allows for calculation of the number of categories, perseverations, and total errors.

The original version of the test (Berg, 1948; Heaton et al., 1993) contained 2 decks of 64 cards and included ambiguous stimuli that could be matched to more than one principle or category, potentially obfuscating interpretation of the deficits and increasing participant frustration. In order to address these problems, the M-WCST removed ambiguous stimuli and shortened the amount of response cards to minimize participant fatigue. The M-WCST has demonstrated high specificity (98.7%) and good sensitivity (45.6%) to brain lesions (Van den Broek et al., 1993) and was able to adequately discriminate between Korsakoff patients and controls (Shoqeirat et al., 1990), as well as between patients with frontal lesions and controls (Zubicaray &

Ashton, 1996). A factor analysis by Nagahama et al. (2003) found that the number of perseverative errors a participant makes represents an aspect of executive dysfunction in Alzheimer`s disease and mild cognitive impairment. The M-WCST also demonstrated robust age effects in hierarchical modeling with controls (Rhodes, 2004).

The Stroop Color and Word Test. The test consists of three pages, each with 100 components randomly organized into five columns. In the first page the participant must read aloud the words “Red”, “Green”, and “Blue” printed in black ink. In the second one, “color naming”, the color (blue, green, or red) of each element “XXXX” must be named. And in the last one, “interference”, the task is to name the color of the ink, inhibiting the reading of the word, which corresponds to the name of another color. The subject has 45 seconds to read aloud, as quickly as possible, the columns from left to right. Finally, the Interference Index was calculated with the formula: $WC - [(W \times C)/(W + C)]$, and indicates the degree to which the person has control over interference (Golden, 2007).

The Stroop Color and Word Test has demonstrated good psychometric properties, (for a review see Strauss et al., 2006), and has been utilized in studies of numerous pathologies including traumatic brain injury (Ben-David et al., 2011; Guise et al., 2014; Ripley et al., 2014), Parkinson’s disease (Sisco et al., 2016), Alzheimer’s disease (Ben-David et al., 2013; Bondi et al., 2002), and stroke (de Bruijn et al., 2014; Rostamian et al., 2015). For this manuscript Total Word-Color and Interference index were considered.

Procedure

The participants completed the M-WCST and Stroop Color and Word Test as part of a large battery of neuropsychological tests that includes the Brief Test of Attention, Rey-Osterrieth Complex Figure test, Phonological and Semantic Verbal

Fluency Test, Boston Naming Test, and Stroop colour and words test, M-WCST, Symbol digit modalities test and Hopkins verbal learning test – Revised (see Guàrdia-Olmos et al., 2015). In order to attain a standard management process of the battery, the following tools and visual aids were established: (a) a randomized list to determine the order of test administration for each participant in order to avoid order bias and cognitive conditioning, (b) a template in Microsoft Excel for entering information to limit bias input information, (c) examples showing the most frequent errors in the administration and scoring of tests, and other tools to maintain standardization in the administration of tests. For further information regarding the study's procedure, see Arango-Lasprilla and Rivera (2015) and Guàrdia-Olmos et al. (2015). The University of Deusto's (Bilbao, Spain) Ethics Committee approved this study.

Statistical Analyses

Demographic variables' effect on neuropsychological performance. The database was randomly divided into two subsamples (A=90% & B=10%). Subsample A (n=4,866) was used to estimate the effect of demographic variables. The effects of demographic variables on M-WCST (Categories, Perseverations, and Total Errors) and Stroop Color and Word Test (Total Word-Color and Interference) scores were evaluated by means of multiple linear regression analyses. The full regression models included the following as predictors: age, age^2, level of education, sex, and all two-way interactions between these variables. Age was centered (= calendar age – mean age in the sample by country) before computing the quadratic age to avoid multicollinearity (Kutner et al., 2005). Education was dummy coded into a variable of 0 and 1: 1 if the participant had > 12 years of education and 0 if the participants had 1-12 years of education (Guàrdia-Olmos et al., 2015), and Sex was dummy coded as male = 1 and female = 0. Independent variables that were not statistically significant in the multiple regression model were removed from the model, and the reduced model was fitted again. In the stepwise model-building procedure, no predictor was removed if it was also included in a higher order term in the model (Aiken et al., 1991). The full regression model can be formally described as: $y_i = \mathrm{B}_0 + \mathrm{B}_1 \cdot (Age - \bar{x}_{Age\ by\ country})_i + \mathrm{B}_2 \cdot \left(Age - \bar{x}_{Age\ by\ country}\right)_i^2 + \mathrm{B}_3 \cdot (Level\ Education)_i + \mathrm{B}_4 \cdot Sex_i + \mathrm{B}_k \cdot\ Interactions_i + \varepsilon_i$. A Bonferroni-corrected alpha-level of .005 (=$.05$ / 10) was used. The model assumes that the residuals ε_i are normally distributed with mean 0 and variance σ_ε^2, i.e., $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$.

Calculation of adjusted Z-score. Adjusted Z-scores for each raw score were calculate using the information provided in each final regression model in a three-step procedure (Rivera & Arango-Lasprilla, 2017; Van Der Elst et al., 2006a, 2006b): 1. The expected test score ($\hat{Y}_i$) is computed based on the fixed effect parameter estimated of the established final regression model: $\hat{Y}_i = B_0 + B_1X_{1i} + B_2X_{2\mathrm{i}} + \cdots + B_{\mathrm{K}}X_{\mathrm{Ki}}$. 2. To obtain the residual value (e_i), a subtraction between the raw score of the neuropsychological test (Y_i) and the predicted value ($\hat{Y}_i$) previously calculated was performed as shown in the following formula: $e_i = Y_i - \hat{Y}_i$. 3. Using the residual standard deviation (SD_e) value provided by the regression model, residuals were standardized: $z_i = e_i\ /SD_e$. This three-step process was applied to each score (M-WCST Categories, M-WCST Perseverations, and M-WCST Total Errors, Stroop Total Word-Color, Stroop Interference) separately for each country. In case scores were not affected by demographic variables, the standardization of the score (z_i) was performed using the formula $z_i =\ x_i - \bar{x}_{sample}/SD_{sample}$.

Multivariate base rates. The exact percentile corresponding to the *Z*-score previously calculated was obtained using the standard normal cumulative distribution function (if the model assumption of normality of the residuals was met in the normative sample), or via the empirical cumulative distribution function of the standardized residuals (if the standardized residuals were not normally distributed in the normative sample). Percentiles that are routinely used in clinical practice or research as indicator of low performance were analyzed in this study: (a) below the 25th percentile, (b) below the 16th percentile, (c) below the 10th percentile, (d) below the 5th percentile, and (e) below the 2nd percentile.

The prevalence below each of these percentiles was calculated. This base-rate analysis was calculated involving examination of executive function performance on the five Z-scores (M-WCST Categories, M-WCST Perseverations, and M-WCST Total Errors; Stroop Total Word-Color, Stroop Interference) simultaneously, not each score in isolation.

Comparative low scores. Subsample B (n=506) was used to estimate the low scores of each participant using the resulting models from subsample A. Then, the low scores were estimated using the normative data from the Stroop test (Golden & Freshwater, 2002; Levine et al., 2004) and the M-WCST (Schretlen, 2010) for the English-speaking population. Finally, a Chi-Square test (X^2) and Wilcoxon Signed-Rank Test were performed to compare the number of low scores obtained using the normative data based on Spanish-speaking population and on the English-speaking population for each specific percentile cutoff groups (25^{th}, 16^{th}, 10^{th}, 5^{th}, and 2^{nd}). Wilcoxon Signed-Rank Test allows to show the differences between the distributions in the low scores without assuming normal distribution. All analyses were performed using SPSS version 23 (IBM Corp., 2015).

Results

Table 2 shows the final regression models for each score (M-WCST Categories, M-WCST Perseverations, and M-WCST Total Errors, Stroop Total Word-Color, Stroop Interference) by country. The amount of variance explained in scores ranged from 3.6% (in El Salvador on the Stroop Interference score) to 43.8% (in Paraguay on the Stroop Total Word-Color). The following scores did not show demographic effects: Stroop Interference [in Honduras (Mean= 1.83; SD= 5.81) and Guatemala (Mean= -1.40; SD= 8.91)], M-WCST Perseverations (In Honduras; Mean= 7.42; SD= 6.50) and M-WCST Total errors (In Honduras; Mean= 15.61; SD= 9.86).

Insert Table 2

The base rates of low test scores on the executive functioning performance are presented in Table 3. Between 49.4% (Puerto Rico) and 77.7% (Argentina) of the sample have at least 1 of the 5 scores below the 25th percentile, and between 32.8% (Puerto Rico) and 48.6% (Guatemala) scored below the 16th percentile on one or more scores. Moreover, between 22.7% (Paraguay) and 34.8% (El Salvador) scored below the 10th percentile on at least 1 of the 5 scores. Between 12.9% (Puerto Rico) and 20.3% (Honduras) scored below the 5th percentile on one or more scores. Finally, between 4.8% (Cuba) and 10.6% (Guatemala) scored below the 2nd percentile on at least one of the five scores.

Insert Table 3

Wilcoxon Signed-Rank Tests showed significant differences between low-score distributions depending on whether English-Speaking or Spanish-speaking normative data were used for their estimation (*p's*<.001; see Table 4). Chi-Squares showed that the number of low scores (No low score; one or more, two or more, three or more low scores, Four or more low scores, five low score) at each specific percentile cutoff group (25^{th}, 16^{th}, 10^{th}, 5^{th}, and 2^{nd}) varied depending on the normative data used to obtain them. For example, using normative data for a Spanish-speaking population, 56.4% of the sample obtained at least 1 low score, while using normative data for an English-Speaking population raised that proportion to 100%. Table 4 shows the possible over-estimation of number of low scores when using normative data from an English-Speaking population in a Spanish-speaking sample.

Insert Table 4

Discussion

Measurement of executive functioning is one of the core domains of a neuropsychological assessment. Executive functioning represents a set of abilities that are necessary for appropriate behavior and independent functioning and can be negatively impacted across the lifespan in neurodevelopmental disorders (e.g. O'Hearn et al., 2008), acquired brain injuries (e.g. Rabinowitz, & Levin, 2014), and dementias (e.g. McKinlay et al., 2010). Therefore, accurate measurement of this domain is imperative and has provided the impetus for Latino-specific normative samples to minimize the likelihood of false positives by neuropsychological assessments.

Considering the multivariate base rates of low scores is another well-established psychometric method for reducing false positives when interpreting test performance (Brooks et al., 2013). When simultaneously considering performance in this large Latino sample across the five scores generated from the M-WCST (M-WCST

Categories, M-WCST Perseverations, and M-WCST Total Errors) and the Stroop Color and Word Test (Stroop Total Word-Color and Stroop Interference), it was found that approximately half of those healthy Latino adults had one or more scores <1SD. This is certainly a high number and should bring caution to interpreting a single isolated low score as being suggestive of problematic executive functioning. It was not considered 'uncommon (<5% of the sample) until there were 4 or more scores <25th percentile, 3 or more scores <16th or <10th percentiles, 2 or more scores <5th percentile, or 1 or more scores <2nd percentile.

Although there is literature considering the multivariate base rates across broad batteries of tests (Brooks & Iverson, 2010; Brooks et al., 2013; Holdnack et al., 2017), the number of studies specifically considering a battery of executive measures is limited in number and is specific to primarily white samples from the United States. Despite the paucity of literature, the existing studies of multivariate base rates on executive functioning measures are supportive of similar findings to the present Latino sample. Brooks and colleagues considered the base rates of low scores in adults and older adults on the Test of Verbal Conceptualization (TVCF; Brooks et al., 2012), a four-subtest battery of executive measures (i.e., Category Fluency, Letter Naming, Classification, and Trails C). When simultaneously considering all five age-adjusted scores from the TVCF (i.e., Category Fluency Total Correct, Letter Naming Total Correct, Classification–Number of Items Correct, Classification–Number of Perseverative Errors, and Trails C–Total Time), they reported that more than one-quarter of adults and more than one-third of older adults had one or more scores <1SD. There were modest differences across education level (standard scores were not adjusted for years education), where those adults and older adults with fewer years of formal schooling had higher rates of low scores.

Multivariate base rates of low scores on the Delis-Kaplan Executive Function System (DKEFS; Delis et al., 2001) have also been examined previously (Crawford et al., 2011; Karr et al., 2017; 2018). Regardless of whether the full DKEFS battery (nine subtests) or an abbreviated four-subtest battery were considered, having low scores across these executive functioning measures was very common (Karr et al., 2017; 2018). For example, having one or more scores <1SD was found in 83% of the sample for the full nine DKEFS subtests (16 total achievement scores) and in 72% of the sample when considering only four subtests (nine total achievement scores). Consistent with a considerable amount of prior literature, the rates of low scores increased with lower intellectual abilities and fewer years of education. For example, nearly 100% of healthy adults and older adults with low average IQ scores or no more than 8 years of education had one or more scores <1SD when given the full DKEFS. In fact, in those with lower intelligence and/or the fewest years of education, it was not considered 'uncommon' (i.e., occurring in fewer than 5% of the sample) until there were 11 or more scores <1SD, 10 or more scores ≤9th percentile, 8 or more scores ≤5th percentile, or 6 or more scores ≤2nd percentile. Clearly these findings have implications for the interpretation of multiple test scores in the domain of executive functioning, but even greater caution is needed when assessing and interpreting test performance in those who fall outside the typical range of functioning for the average English-speaking North American.

An example is provided to facilitate comprehension of the use of low scores and its clinical utility. Calculation of low scores may be tedious and increase the chance for making mistakes due to the multiple computations that need to be conducted. Therefore, authors created a calculator in Microsoft Excel in which the clinician must include the following information: country, age, education, sex, and raw scores, to calculate the z score, percentile, and number of low scores. This tool is freely available for all users and can be downloaded at https://neuropsychologylearning.com/datos-normativos-archivos-descargables/

Example: A Colombian man of 55-year-old with 13 years of education obtained the following scores: Stroop Word-Color = 30, Stroop Interference = -5, M-WCST Categories = 1, M-WCST Perseverations = 11, and M-WCST Total errors = 15. Once the number of low scores is obtained using the calculator (see Figure 1), Table 3 should be consulted. Suppose you are interested in the number of low scores below the 10^{th} percentile. In this case, this person obtained two low scores below the 10^{th} percentile, indicating that he belongs to the 12.4% of the Colombian sample who obtained the same result.

Insert Figure 1.

This study highlights the importance of using appropriate normative data for each population. Using normative data from a different population can lead to over or underestimation of low scores there by increasing the probability of coming up to an erroneous diagnosis of cognitive deficit. There are several differences between countries and cultures that may explain these low-score distribution variances, such as socioeconomic status, health-system, nutrition and safety, literacy, quality of education, familiarity with testing environment, among others (Rivera et al., 2019).

These results should be interpreted in the light of the following limitations: (a) In this study, the MVBRs were calculated for two of the most commonly used tests to measure executive functioning processes; however, we do not know if these results are similar or different when other tests are used. (b) The number of tests used in the present study was five, and thus it is possible that to the extent that more scores from other executive functioning tests were included, these results could even be lower. (c) The present study was conducted with a large Spanish-speaking population from 11 Latin American countries and Puerto Rico, and for this reason it is not possible to generalize these results to those countries outside of the present sample or those whose language is not Spanish (e.g., Brazil). (d) It is possible that the low scores found in this study could be explained by some variables that were not measured or not considered when carrying out the study, such level of bilingualism and the quality of education, among others. (e) Education was used as a dummy coded, dichotomous variable (i.e., 12 or > 12 years of education), and as such, future studies should include education as a continuous variable. Finally, (f) the sample was not stratified by intellectual level, which has been shown to be associated with different base rates of low scores on cognitive measures (Brooks et al., 2013; Brooks et al., 2011; Brooks et al., 2009; Brooks et al., 2008; Guàrdia-Olmos et al., 2015; Rivera & Arango-Lasprilla, 2017). Future research will consider MVBRs in Latino samples with varying levels of intellectual abilities.

Reducing misdiagnosis of executive dysfunction through appropriate normative samples and multivariate base rate interpretation will advance neuropsychological assessments with Spanish-speaking adults from Argentina, Bolivia, Chile, Colombia, Cuba, El Salvador, Guatemala, Honduras, Mexico, Paraguay, Peru, and Puerto Rico. The results of this study are similar to other studies with the TVCF and DKEFS in North American samples, which provides support that interpretation of isolated low scores in the absence of knowledge about multivariate base rates can lead to spuriously high rates of executive impairment. We support using multivariate base rates, in collaboration with clinical judgment, other findings, and relevant history, as an appropriate clinical tool. The wide differences in the number of low scores obtained when using normative data from countries or cultures different from those of the patient highlight the need to develop, using updated methodologies, normative data specific to each country or culture.

Conflict of interest details

Dr. Brooks reports the following conflicts of interest: co-author of the Child and Adolescent Memory Profile (ChAMP, Sherman and Brooks, 2015, PAR Inc.), Memory Validity Profile (MVP; Sherman and Brooks, 2015, PAR Inc.), and Multidimensional Everyday Memory Ratings for Youth (MEMRY, Sherman and Brooks, 2017, PAR Inc.), and he receives royalties for the sales of these tests; co-editor of the Pediatric Forensic Neuropsychology textbook (2012, Oxford University Press) and receives royalties for the sales of this book; previously been provided with free test credits from CNS Vital Signs as an in-kind support for his research.

Funding details

Brian Brooks acknowledges partial salary funding from the Canadian Institutes for Health Research (CIHR) Embedded Clinician Researcher Salary Award.

References

Aiken, L. S., West, S. G., & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage Publications, Inc.

Arango-Lasprilla, J. C., Rivera, D., Longoni, M., Saracho, C. P., Garza, M. T., Aliaga, A., Rodríguez, W., Rodríguez-Agudelo, Y., Rábago, B., Sutter, M., Schebela, S. Luna, M., Ocampo-Barba, N., Galarza-del-Angel, J., Bringas, M.L., Esenarro, L., Martínez, C., García-Egan, P., & Perrin, P.B. (2015). Modified Wisconsin Card Sorting Test (M-WCST): Normative data for the Latin American Spanish speaking adult population. *NeuroRehabilitation*, *37*(4), 563–590. https://doi.org/10.3233/NRE-151280

Arango-Lasprilla, J., & Rivera, D. (2015). *Neuropsicología en Colombia: Datos normativos, estado actual y retos a futuro* [Neuropsychology in Colombia: Normative data, current state and future challenges]. Editorial Universidad Autónoma de Manizales.

Baron, I.S. (2018). *Neuropsychological Evaluation of the Child (2nd Edition).* Oxford University Press.

Ben-David, B.M., Nguyen, L.L.T., & van Lieshout, P.H.H.M. (2011). Stroop Effects in Persons with Traumatic Brain Injury: Selective Attention, Speed of Processing, or Color-Naming? A Meta-analysis. *Journal of the International Neuropsychological Society*, *17*(02), 354–363. https://doi.org/10.1017/S135561771000175X

Ben-David, B.M., Tewari, A., Shakuf, V., & Van Lieshout, P.H.H.M. (2013). Stroop Effects in Alzheimer's Disease: Selective Attention Speed of Processing, or Color-naming? A Meta-Analysis. *Journal of Alzheimer's Disease*, *38*(4), 923–938. https://doi.org/10.3233/JAD-131244

Berg, E. A. (1948). A Simple Objective Technique for Measuring Flexibility in Thinking. *The Journal of General Psychology*, *39*(1), 15–22. https://doi.org/10.1080/00221309.1948.9918159

Binder, L. M., Iverson, G. L., & Brooks, B. L. (2009). To err is human: "abnormal" Neuropsychological scores and variability are common in healthy adults. *Archives of Clinical Neuropsychology*, *24*(1), 31–46. https://doi.org/10.1093/arclin/acn001

Bondi, M. W., Monsch, A. U., Butters, N., Salmon, D. P., & Paulsen, J. S. (1993). Utility of a modified version of the wisconsin card sorting test in the detection of dementia of the alzheimer type. *Clinical Neuropsychologist*, *7*(2), 161–170. https://doi.org/10.1080/13854049308401518

Bondi, M. W., Serody, A. B., Chan, A. S., Eberson-Shumate, S. C., Delis, D. C., Hansen, L. A., & Salmon, D. P. (2002). Cognitive and neuropathologic correlates of Stroop Color-Word Test performance in Alzheimer's disease. *Neuropsychology*, *16*(3), 335–343. http://dx.doi.org/10.1037/0894-4105.16.3.335

Borkowska, A., Drozdz, W., Jurkowski, P., & Rybakowski, K. (2009). The Wisconsin Card Sorting Test and the N-back test in mild cognitive impairment and elderly depression. *The World Journal of Biological Psychiatry*, *10*(4), 870 -876. https://doi.org/10.1080/15622970701557985

Brooks, B. L., & Iverson, G. L. (2010). Comparing actual to estimated base rates of "abnormal" scores on neuropsychological test batteries: Implications for interpretation. *Archives of Clinical Neuropsychology*, *25*(1), 14–21. https://doi.org/10.1093/arclin/acp100

Brooks, B. L., Iverson, G. L., & Holdnack, J. A. (2013). Understanding and using multivariate base rates with the WAIS–IV/WMS–IV. In *WAIS-IV, WMS-IV, and ACS* (pp. 75-102). Academic Press.

Brooks, B.L., Iverson, G.L., Koushik, N.S., Mazur-Mosiewicz, A., Horton Jr, A.M., & Reynolds, C. R. (2013). Prevalence of low scores in children and adolescents on the test of verbal conceptualization and fluency. *Applied Neuropsychology: Child*, *2*(1), 70-77. https://doi.org/10.1080/21622965.2012.742792

Brooks, B.L., Iverson, G.L., Lanting, S.C., Horton, A.M., & Reynolds, C.R. (2012). Improving Test Interpretation for Detecting Executive Dysfunction in Adults and Older Adults: Prevalence of Low Scores on the Test of Verbal Conceptualization and Fluency. *Applied Neuropsychology*, *19*(1), 61–70. https://doi.org/10.1080/09084282.2012.651951

Brugnolo, A., De Carli, F., Accardo, J., Amore, M., Bosia, L. E., Bruzzaniti, C., Cappa, S.F., Cocit, L., Colazzo, G., Ferrara, M., Ghio, L., Magi, E., Mancardi, G.L., Nobili, F., Pardini, M., Rissotto, R., Serrati, C., & Girtler, N. (2016). An updated Italian normative dataset for the Stroop color word test (SCWT). *Neurological Sciences*, *37*(3), 365–372. https://doi.org/10.1007/s10072-015-2428-2

Cianchetti, C., Corona, S., Foscoliano, M., Contu, D., & Sannio-Fancello, G. (2007). Modified Wisconsin Card Sorting Test (MCST, MWCST): Normative Data in Children 4–13 Years Old, According to Classical and New Types of Scoring. *The Clinical Neuropsychologist*, *21*(3), 456–478. https://doi.org/10.1080/13854040600629766

Cook, N. E., Karr, J. E., Brooks, B. L., Garcia-Barrera, M. A., Holdnack, J. A., & Iverson, G. L. (2019). Multivariate base rates for the assessment of executive functioning among children and adolescents. *Child Neuropsychology*, *25*(6), 836-858. https://doi.org/10.1080/09297049.2018.1543389

Crawford, J. R., Garthwaite, P. H., Sutherland, D., & Borland, N. (2011). Some supplementary methods for the analysis of the Delis–Kaplan Executive Function System. *Psychological assessment,* *23*(4), 888-898. http://dx.doi.org/10.1037/a0023712

de Bruijn, M. A. A. M., Synhaeve, N. E., van Rijsbergen, M. W. A., de Leeuw, F.-E., Jansen, B. P. W., & de Kort, P. L. M. (2014). Long-Term Cognitive Outcome of Ischaemic Stroke in Young Adults. *Cerebrovascular Diseases*, *37*(5), 376–381. https://doi.org/10.1159/000362592

del Pino, R., Peña, J., Ibarretxe-Bilbao, N., Schretlen, D. J., & Ojeda, N. (2016). [Modified Wisconsin Card Sorting Test: standardization and norms of the test for a population sample in Spain]. *Revista de Neurologia*, *62*(5), 193–202.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan executive function system (D–KEFS)* [Database record]. APA PsycTests. https://doi.org/10.1037/t15082-000

Diamond, A. (2013). Executive functions. *Annual review of psychology*, *64*, 135-168. https://doi.org/10.1146/annurev-psych-113011-143750

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186-204. https://doi.org/10.1016/j.cortex.2016.04.023

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Minimental state”: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189-198. https://doi.org/10.1016/0022-3956(75)90026-6

Fossati, P., Ergis, A. M., & Allilaire, J. F. (2001). Problem-solving abilities in unipolar depressed patients: Comparison of performance on the modified version of the

Wisconsin and the California sorting tests. *Psychiatry Research*, *104*(2), 145–156. https://doi.org/10.1016/S0165-1781(01)00307-9

Golden C. J. (2007). *Stroop, Test de Colores y Palabras: Manual* (5º Edición). Madrid: Tea Ediciones, S.A.

Golden, C., & Freshwater, S. M. (2002). *Stroop Color and Word Test. Revised 2002 adult manual for clinical and Experimental uses*. Wood Dale, Illinois: Stoelting.

Gotham, A. M., Brown, R. G., & Marsden, C. D. (1988). “Frontal” cognitive function in patients with Parkinson’s disease “on” and “off” levodopa. *Brain : A Journal of Neurology*, *111*(2), 299–321. https://doi.org/10.1093/brain/111.2.299

Guàrdia-Olmos, J., Peró-Cebollero, M., Rivera, D., & Arango-Lasprilla, J. C. (2015). Methodology for the development of normative data for ten Spanish-language neuropsychological tests in eleven Latin American countries. *NeuroRehabilitation, 37*(4), 493-499. https://doi.org/10.3233/nre-151277

Guise, B. J., Thompson, M. D., Greve, K. W., Bianchini, K. J., & West, L. (2014). Assessment of performance validity in the Stroop Color and Word Test in mild traumatic brain injury patients: A criterion-groups validation design. *Journal of Neuropsychology*, *8*(1), 20–33. https://doi.org/10.1111/jnp.12002

Hartman, M., Steketee, M., & Silva, S. (2003). Wisconsin Card Sorting Test performance in schizophrenia: the role of working memory. *Schizophrenia research, 63*(3), 201–17.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test (WCST): Manual: Revised and Expanded*. Psychological Assessment Resources.

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences*, *16*(3), 174–180.

https://doi.org/10.1016/j.tics.2012.01.006

Holdnack, J. A., Tulsky, D. S., Brooks, B. L., Slotkin, J., Gershon, R., Heinemann, A. W., & Iverson, G. L. (2017). Interpreting Patterns of Low Scores on the NIH Toolbox Cognition Battery. *Archives of Clinical Neuropsychology*, *32*(5), 574–584. https://doi.org/10.1093/arclin/acx032

Horton, A.M. Jr, & Wedding, D. (2008). *The Neuropsychology Handbook, Third Edition*. Springer.

IBM Corp. (2015). *SPSS Version 23.0.* Armonk, NY: IBM Corp.

Ikeda, Y., Okuzumi, H., Kokubun, M., & Haishi, K. (2011). Age-Related Trends of Interference Control in School-Age Children and Young Adults in the Stroop Color–Word Test. *Psychological Reports*, *108*(2), 577–584. https://doi.org/10.2466/04.10.22.PR0.108.2.577-584

Ingraham, L. J. (1996). An empirical approach to determining criteria for abnormality in test batteries with multiple measures. *Neuropsychology*, *10*(1), 120–124. https://doi.org/10.1037/0894-4105.10.1.120

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: a review of our current understanding. *Neuropsychology review, 17*(3), 213-233. https://doi.org/10.1007/s11065-007-9040-z

Kado, Y., Sanada, S., Yanagihara, M., Ogino, T., Abiru, K., & Nakano, K. (2004). Effect of development and aging on the modified Wisconsin Card Sorting Test in normal subjects. *No to Hattatsu = Brain and Development*, *36*(6), 475–480.

Karr, J. E., Garcia-Barrera, M. A., Holdnack, J. A., & Iverson, G. L. (2016). Using Multivariate Base Rates to Interpret Low Scores on an Abbreviated Battery of the Delis–Kaplan Executive Function System. *Archives of Clinical Neuropsychology, 32*(3), 297-305. https://doi.org/10.1093/arclin/acw105

Karr, J. E., Garcia-Barrera, M. A., Holdnack, J. A., & Iverson, G. L. (2017). Using Multivariate Base Rates to Interpret Low Scores on an Abbreviated Battery of the Delis–Kaplan Executive Function System. *Archives of Clinical Neuropsychology*, *32*(3), 297–305. https://doi.org/10.1093/arclin/acw105

Karr, J. E., Garcia-Barrera, M. A., Holdnack, J. A., & Iverson, G. L. (2018). Advanced clinical interpretation of the Delis-Kaplan Executive Function System: multivariate base rates of low scores. *The Clinical Neuropsychologist, 32*(1), 42-53. https://doi.org/10.1080/13854046.2017.1334828

Kolb, B. & Whishaw, I.Q. (2015). *Fundamentals of Human Neuropsychology (7th Edition)*. Macmillan Learning.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ- 9. *Journal of General Internal Medicine, 16*(9), 606-613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models (5th ed.)*. McGraw Hill.

Lezak, M. D. (2012a). *Neuropsychological assessment*. Oxford University Press.

Lezak, M. D. (2012b). *Neuropsychological assessment* (5th ed.). Oxford University Press.

Levine, A.J., Miller, E.N., Becker, J.T., Selnes, O.A., & Cohen, B.A. (2004). Normative data for determining significance of test–retest differences on eight common neuropsychological instruments. *The Clinical Neuropsychologist*, *18*(3), 373-384. http://dx.doi.org/10.1080/1385404049052420

Llinas-Regla, J., Vilalta-Franch, J., Lopez-Pousa, S., Calvo-Perxas, L., & Garre-Olmo, J. (2013). Demographically Adjusted Norms for Catalan Older Adultson the Stroop Color and Word Test. *Archives of Clinical Neuropsychology*, *28*(3), 282–

296. https://doi.org/10.1093/arclin/act003

Mahoney, F. I., & Barthel, D. (1965). Functional evaluation: The Barthel Index. *Maryland State Medical Journal, 14,* 56-61.

Maqbool, M., Sinha, V. K., & Vikas. (2015). Sustained Attention and Executive Functioning among Remitted Adolescents with Bipolar Disorder. *Indian Journal of Psychological Medicine*, *37*(3), 322–6. https://doi.org/10.4103/0253-7176.162955

McKinlay, A., Grace, R. C., Dalrymple-Alford, J. C., & Roger, D. (2010). Characteristics of executive function impairment in Parkinson's disease patients without dementia. *Journal of the International Neuropsychological Society*, *16*(2), 268-277.

Morgan, J.E. and Ricker, J.H. (2018). *Textbook of Clinical Neuropsychology (2nd Edition)*. Routledge.

Nagahama, Y., Okina, T., Suzuki, N., Matsuzaki, S., Yamauchi, H., Nabatame, H., & Matsuda, M. (2003). Factor Structure of a Modified Version of the Wisconsin Card Sorting Test: An Analysis of Executive Deficit in Alzheimer's Disease and Mild Cognitive Impairment. *Original Research Article Dement Geriatr Cogn Disord*, *16*(2), 103–112. https://doi.org/10.1159/000070683

Nakayama, H., Osawa, M., & Maruyama, S. (1990). [Aging effect on frontal lobe function-evaluation by the new modified Wisconsin Card Sorting Test]. *No to Shinkei = Brain and Nerve*, *42*(8), 765–71.

Nelson, H. E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *12*(4), 313–24.

O'Hearn, K., Asato, M., Ordaz, S., & Luna, B. (2008). Neurodevelopment and executive function in autism. *Development and psychopathology*, *20*(4), 1103-1132. https://doi.org/10.1017/S0954579408000527

Olabarrieta-Landa, L., Ramos Usuga, D., Rivera, D., Leal, G., Bailey, K. C., Calderón Chagualá, A., Rabago, B., Esenarro, L., Mascialino, G., & Arango-Lasprilla, J. C. (2019). Prevalence of low scores on language tests as a potential factor in misdiagnosis of cognitive impairment in a Spanish-speaking adult population. *Applied Neuropsychology: Adult*, 1-12. https://doi.org/10.1080/23279095.2019.1704409

Paulo, A. C., Sampaio, A., Santos, N. C., Costa, P. S., Cunha, P., Zihl, J., … & Sousa, N. (2011). Patterns of Cognitive Performance in Healthy Ageing in Northern Portugal: A Cross-Sectional Analysis. *PLoS ONE*, *6*(9), e24553. https://doi.org/10.1371/journal.pone.0024553

Peinemann, A., Schuller, S., Pohl, C., Jahn, T., Weindl, A., & Kassubek, J. (2005). Executive dysfunction in early stages of Huntington's disease is associated with striatal and insular atrophy: A neuropsychological and voxel-based morphometric study. *Journal of the Neurological Sciences*, *239*(1), 11–19. https://doi.org/10.1016/J.JNS.2005.07.007

Peña-Casanova, J., Quiñones-Úbeda, S., Gramunt-Fombuena, N., Quintana, M., Aguilar, M., Molinuevo, J. L., Serradell, M., Robles, A., Barquero, M.S., Payno, M., Antúnez, C., Martínez-Parra, C., Frank-García, A., Fernández, M., Alfonso, V., Sol, J.M., & Blesa, R. (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(4), 413-429.

Rabin, L. A., Paolillo, E., & Barr, W. B. (2016). Stability in test-usage practices of clinical neuropsychologists in the United States and Canada over a 10-year period: A follow-up survey of INS and NAN members. *Archives of Clinical Neuropsychology, 31*(3), 206-230. https://doi.org/10.1093/arclin/acw007

Rabinowitz, A. R., & Levin, H. S. (2014). Cognitive sequelae of traumatic brain injury. *The Psychiatric Clinics of North America*, *37*(1), 1- 11 https://doi.org/10.1016/j.psc.2013.11.004

Rhodes, M. G. (2004). Age-Related Differences in Performance on the Wisconsin Card Sorting Test: A Meta-Analytic Review. *Psychology and Aging*, *19*(3), 482–494. https://doi.org/10.1037/0882-7974.19.3.482

Ripley, D. L., Morey, C. E., Gerber, D., Harrison-Felix, C., Brenner, L. A., Pretz, C. R., Cusick, C., & Wesnes, K. (2014). Atomoxetine for attention deficits following traumatic brain injury: Results from a randomized controlled trial. *Brain Injury*, *28*(12), 1514–1522. https://doi.org/10.3109/02699052.2014.919530

Rivera, D., Olabarrieta-Landa, L., Brooks, B.L., Ertl, M.M., Benito-Sánchez, I., Quijano, M.C., Rodriguez-Irizarry, W., Aguayo Arelis, A., Rodríguez-Agudelo, Y., & Arango-Lasprilla, J.C. (2019). Multivariate Base Rates of Low Scores on Tests of Learning and Memory Among Latino Adult Populations. *Journal of the International Neuropsychological Society: JINS*, *25*(8), 834-844. https://doi.org/10.1017/S135561771900050X

Rivera, D., & Arango-Lasprilla, J. C. (2017). Methodology for the development of normative data for Spanish-speaking pediatric populations. *NeuroRehabilitation*, *41*(3), 581-592. https://doi.org/10.3233/nre-172275

Rivera, D., Perrin, P. B., Stevens, L. F., Garza, M. T., Weil, C., Saracho, C. P., Rodíguez, W., Rodriguez-Agudelo, Y., Rábago, B., Weiler, G., García de la

Cadena, C., Longoni, M., Martínez, C., Ocampo-Barba, N., Aliaga, A., Galarza-del-Angel, J., Guerra, A., Esenarro, L., & Arango-Lasprilla, J. C. (2015). Stroop Color-Word Interference Test: Normative data for the Latin American Spanish speaking adult population. *NeuroRehabilitation*, *37*(4), 591–624. https://doi.org/10.3233/NRE-151281

Rivera, D., Olabarrieta-Landa, L., Van der Elst, W., Gonzalez, I., Rodríguez-Agudelo, Y., Aguayo Arelis, A., Rodriguez-Irizarry, W., García de la Cadena, C., & Arango-Lasprilla, J. C. (2019). Normative data for verbal fluency in healthy Latin American adults: Letter M, and fruits and occupations categories. *Neuropsychology*, *33*(3), 287-300. https://doi.org/10.1037/neu0000518

Rognoni, T., Casals-Coll, M., Sánchez-Benavides, G., Quintana, M., Manero, R. M., Calvo, L., Palomo, R., Aranciva, F., Tamayo, F., & 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. *Neurologia*, *28*(2), 73–80. https://doi.org/10.1016/j.nrl.2012.02.009

Rostamian, S., van Buchem, M. A., Westendorp, R. G. J., Jukema, J. W., Mooijaart, S. P., Sabayan, B., & de Craen, A. J. M. (2015). Executive function, but not memory, associates with incident coronary heart disease and stroke. *Neurology*, *85*(9), 783–789. https://doi.org/10.1212/WNL.0000000000001895

Savitz, J. B., & Jansen, P. (2003). The Stroop Color-Word Interference Test as an Indicator of ADHD in Poor Readers. *The Journal of Genetic Psychology*, *164*(3), 319–333. https://doi.org/10.1080/00221320309597986

Schretlen, D. J. (2010). *M-WCST (Modified Wisconsin Card Sorting Test)*. Psychological Assessment Resources.

Schretlen, D. J., Testas, S. M., Winicki, J. M., Pearlson, G. D., & Gordon, B. (2008). Frequency and bases of abnormal performance by healthy adults on neuropsychological testing. *Journal of the International Neuropsychological Society*, *14*(3), 436–445. https://doi.org/10.1017/S1355617708080387

Shoqeirat, M. A., Mayes, A., MacDonald, C., Meudell, P., & Pickering, A. (1990). Performance on tests sensitive to frontal lobe lesions by patients with organic amnesia: Leng & Parkin revisited. *British Journal of Clinical Psychology*, *29*(4), 401–408. https://doi.org/10.1111/j.2044-8260.1990.tb00903.x

Sisco, S. M., Slonena, E., Okun, M. S., Bowers, D., & Price, C. C. (2016). Parkinson's disease and the Stroop color word test: processing speed and interference algorithms. *The Clinical Neuropsychologist*, *30*(7), 1104–1117. https://doi.org/10.1080/13854046.2016.1188989

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd Editio). New York, NY: Oxford University Press.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Stuss, D. T. (2011). Functions of the frontal lobes: Relation to executive functions. *Journal of the International Neuropsychological Society, 17*(5), 759. https://doi.org/10.1017/S1355617711000695

Van den Broek, M. D., Bradshaw, C. M., & Szabadi, E. (1993). Utility of the Modified Wisconsin Card Sorting Test in neuropsychological assessment. *British Journal of Clinical Psychology*, *32*(3), 333–343. https://doi.org/10.1111/j.2044-8260.1993.tb01064.x

Van Der Elst, W., Van Boxtel, M. P., Van Breukelen, G. J., & Jolles, J. (2006a). Normative data for the Animal, Profession and Letter M Naming verbal fluency tests for Dutch speaking participants and the effects of age, education, and sex. *Journal of the International Neuropsychological Society*, *12*(01), 80-89. https://doi.org/10.1017/s1355617706060115

Van Der Elst, W., Van Boxtel, M. P., Van Breukelen, G. J., & Jolles, J. (2006b). The Letter Digit Substitution Test: Normative data for 1,858 healthy participants aged 24–81 from the Maastricht Aging Study (MAAS): influence of age, education, and sex. *Journal of Clinical and Experimental Neuropsychology*, *28*(6), 998-1009. https://doi.org/10.1080/13803390591004428

Villaseñor-Cabrera, T., Guàrdia-Olmos, J., Jiménez-Maldonado, M., Rizo-Curiel, G., & Peró-Cebollero, M. (2010). Sensitivity and specificity of the Mini-Mental State Examination in the Mexican population. *Quality & Quantity, 44*(6), 1105-1112. https://doi.org/10.1007/s11135-009-9263-6

Yeates, K.O., Ris, M.D., Taylor, H.G., & Pennington, B.F. (2010). *Pediatric Neuropsychology: Research, Theory, and Practice*. The Guilford Press.

Zimmermann, N., Cardoso, C. de O., Trentini, C. M., Grassi-Oliveira, R., & Fonseca, R. P. (2015). Brazilian preliminary norms and investigation of age and education effects on the Modified Wisconsin Card Sorting Test, Stroop Color and Word test and Digit Span test in adults. *Dementia & Neuropsychologia*, *9*(2), 120–127. https://doi.org/10.1590/1980-57642015DN92000006

Zubicaray, G. de, & Ashton, R. (1996). Nelson's (1976) modified card sorting test: A review. *The Clinical Neuropsychologist*, *10*(3), 245–254. https://doi.org/10.1080/13854049608406687

Table 1.

Sample distribution by country, age, education, and sex.

	n	**Age**	**Education**	**Sex**	
				Male	**Female**
		***Mean* (SD)**	**Mean (SD)**	***n* (%)**	***n* (%)**
Argentina	320	45.7 (19.5)	13.8 (4.5)	96 (30.0%)	224 (70.0%)
Bolivia	274	55.8 (22.0)	8.5 (4.4)	99 (36.1%)	175 (63.9%)
Chile	320	55.1 (19.6)	10.0 (5.2)	134 (41.9%)	186 (58.1%)
Colombia	1425	58.2 (19.6)	9.6 (5.3)	610 (42.8%)	815 (57.2%)
Cuba	306	53.0 (19.7)	11.7 (3.7)	142 (46.4%)	164 (53.6%)
El Salvador	257	56.0 (20.7)	8.9 (5.3)	100 (38.9%)	157 (61.1%)
Guatemala	214	53.2 (17.4)	11.5 (5.7)	95 (44.4%)	119 (55.6%)
Honduras	184	48.6 (18.8)	8.6 (5.6)	67 (36.4%)	117 (63.6%)
Mexico	1270	52.5 (20.5)	9.3 (4.7)	422 (33.2%)	848 (66.8%)
Paraguay	263	53.0 (14.8)	9.5 (4.4)	101 (38.4%)	162 (61.6%)
Peru	245	43.4 (20.6)	14.1 (3.7)	87 (35.5%)	158 (64.5%)
Puerto Rico	294	50.9 (18.5)	13.2 (4.2)	126 (42.9%)	168 (57.1%)

Table 2.

Beta coefficients and R^2 for each score and country (n=4,866).

Score		Argentina	Bolivia	Chile	Colombia	Cuba	El Salvador	Guatemala	Honduras	Mexico	Paraguay	Peru	Puerto Rico
Stroop Total Word-Color	Intercept	38.839	32.252	29.993	32.167	35.167	27.094	32.399	28.749	34.818	28.242	39.100	37.795
	Age	-0.223	-0.349	-0.354	-0.417	-0.422	-0.212	-0.251	-0.209	-0.355	-0.114	-0.337	-0.465
	Age^2	-0.006	--	--	-0.003	--	--	-0.008	--	--	--	--	--
	Education	4.985	--	11.266	4.910	3.106	10.765	6.135	5.795	4.264	10.536	--	--
	Sex	--	--	--	--	--	--	--	--	2.481	--	--	--
	Age x Edu	--	--	--	--	0.397	--	--	--	--	-0.293	--	--
	R^2	.341	.215	.418	.371	.246	.269	.219	.226	.304	.438	.272	.376
Stroop Interference	Intercept	-1.581	3.009	-1.245	-1.478	0.852	0.581	--	--	0.581	-4.648	0.174	1.626
	Age	-0.120	--	-0.113	-0.090	-0.086	-0.079	--	--	-0.130	-0.001	-0.151	-0.212
	Age^2	--	--	--	--	--	--	--	--	--	--	--	--
	Education	2.681	--	4.551	--	-1.723	--	--	--	--	3.636	--	--
	Sex	--	--	--	2.768	--	--	--	--	2.401	--	--	--
	Age x Edu	--	--	--	--	0.249	--	--	--	--	-0.316	--	--
	R^2	.142	--	.137	.071	.044	.036	--	--	.080	.197	.098	.134
M-WCST Categories	Intercept	5.443	4.369	5.235	3.438	4.496	3.275	3.950	3.109	4.572	5.054	3.778	5.155
	Age	--	-0.024	-0.018	-0.030	-0.024	-0.018	--	-0.040	-0.027	-0.020	-0.032	-0.030
	Age^2	--	--	--	--	--	--	--	--	-4E-03	-0.001	--	--
	Education	0.395	--	--	1.035	0.702	2.090	1.406	1.048	0.712	0.658	1.189	--
	Sex	--	--	--	--	--	--	--	--	--	--	--	--
	R^2	.050	.089	.070	.195	.138	.183	.142	.208	.130	.224	.329	.140

M-WCST Perseverations	Intercept	5.115	7.072	3.048	8.262	5.191	7.994	7.914	--	4.707	7.319	3.756	4.596
	Age	0.037	0.103	0.039	0.088	0.083	0.066	-0.033	--	0.107	0.062	0.046	0.086
	Age^2	--	--	--	--	--	--	-0.003	--	0.001	--	0.002	--
	Education	-2.491	--	--	-3.141	--	-5.214	-6.246	--	-2.088	-2.705	-2.152	-2.421
	Sex	--	--	-1.568	--	--	--	--	--	--	-1.084	2.406	-
	Age x Edu	--	--	--	--	--	--	0.237	--	--	--		--
	Age^2 x Edu	--	--	--	--	--	--	0.011	--	--	--	--	--
	Age x Sex											0.112	
	R^2	.112	.099	.058	.110	.056	.153	.137	--	.124	.350	.350	.109
M-WCST Total Errors	Intercept	9.177	14.566	8.847	17.394	13.481	18.956	14.818	--	13.181	13.148	15.016	9.275
	Age	0.054	0.146	0.115	0.131	0.136	0.103	--	--	0.176	0.191	0.190	0.198
	Age^2	--	--	--	--	--	--	--	--	--	0.007	--	--
	Education	-4.030	--	--	-5.413	-4.388	-11.924	-7.106	--	-4.041	-4.556	-4.659	--
	Sex	--	--	-2.739	--	--	--	--	--	--	-2.329	--	--
	R^2	.149	.082	.100	.151	.113	.217	.104	--	.142	.414	.254	.138

Table 3.

Cumulative percent with the specified number of adjusted executive function low scores below the specified percentile cutoff (n=4,866).

	Number of low scores	Argentina	Bolivia	Chile	Colombia	Cuba	El Salvador	Guatemala	Honduras	Mexico	Paraguay	Peru	Puerto Rico
<25th percentile	None	22.4%	40.7%	42.9%	41.2%	44.4%	35.2%	37.6%	37.3%	45.5%	40.9%	48.2%	50.6%
	1+	77.7%	59.3%	57.1%	58.9%	55.6%	64.7%	62.4%	62.7%	54.6%	59.2%	51.9%	49.4%
	2+	42.9%	35.4%	38.6%	37.4%	34.1%	31.4%	39.1%	35.9%	36.6%	34.4%	34.9%	28.3%
	3+	23.5%	18.1%	20.4%	17.4%	21.1%	18.1%	20.6%	18.3%	21.2%	15.8%	20.7%	21.9%
	4+	8.4%	6.2%	8.0%	8.4%	7.0%	6.2%	7.9%	4.6%	8.5%	7.1%	10.6%	14.0%
	5	2.0%	2.1%	1.1%	2.6%	2.6%	2.4%	3.7%	1.3%	4.1%	2.1%	3.7%	5.7%
<16th percentile	None	63.2%	59.3%	58.5%	57.9%	63.0%	51.9%	51.3%	51.6%	61.3%	59.9%	61.0%	67.2%
	1+	36.8%	40.8%	41.5%	42.2%	37.0%	48.1%	48.6%	48.4%	38.8%	40.2%	39.0%	32.8%
	2+	19.1%	20.6%	24.0%	22.2%	21.8%	20.0%	23.2%	26.2%	23.1%	16.6%	22.9%	18.5%
	3+	8.4%	9.1%	10.2%	8.9%	11.4%	9.0%	10.5%	11.8%	10.1%	8.7%	10.1%	14.0%
	4+	3.4%	2.5%	0.7%	3.1%	1.8%	1.9%	4.2%	2.0%	3.7%	3.3%	3.7%	9.1%
	5	0.7%	0.4%	0.0%	0.8%	0.7%	0.5%	0.5%	0.0%	1.6%	0.0%	1.4%	3.4%
<10th percentile	None	72.2%	72.8%	73.5%	71.0%	74.1%	65.2%	69.8%	66.7%	72.6%	77.3%	73.9%	75.8%
	1+	27.7%	27.1%	26.6%	29.0%	25.9%	34.8%	30.1%	33.4%	27.3%	22.7%	26.2%	24.1%
	2+	13.3%	15.6%	14.2%	12.4%	10.3%	11.0%	13.2%	13.8%	13.4%	9.9%	12.0%	12.4%
	3+	4.9%	3.3%	5.1%	4.0%	5.9%	4.3%	3.1%	3.3%	5.7%	5.8%	6.0%	9.4%
	4+	1.6%	0.4%	0.0%	0.7%	1.1%	1.0%	0.5%	1.3%	1.6%	0.8%	1.9%	4.9%
	5	0.3%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%	0.5%	0.0%	0.5%	1.9%

$<5^{th}$ percentile	None	84.6%	82.7%	84.4%	83.4%	87.0%	84.3%	82.0%	79.7%	84.3%	86.8%	84.4%	87.2%
	1+	15.4%	17.3%	15.7%	16.8%	13.0%	15.7%	18.0%	20.3%	15.7%	13.2%	15.7%	12.9%
	2+	6.7%	7.0%	6.2%	5.6%	6.0%	4.3%	7.4%	6.6%	6.1%	5.8%	7.4%	7.6%
	3+	1.0%	0.8%	0.7%	1.2%	2.3%	0.5%	1.6%	1.4%	2.0%	2.1%	2.8%	4.6%
	4+	0.0%	0.0%	0.0%	0.3%	0.4%	0.0%	0.5%	0.7%	0.2%	0.0%	0.0%	2.3%
	5	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%
$<2^{th}$ percentile	None	90.0%	91.4%	93.8%	92.8%	95.2%	92.4%	89.4%	91.5%	93.3%	91.7%	93.6%	93.2%
	1+	10.0%	8.7%	6.2%	7.1%	4.8%	7.6%	10.6%	8.5%	6.7%	8.3%	6.5%	6.9%
	2+	4.0%	2.1%	2.2%	1.9%	2.6%	1.9%	3.7%	2.0%	1.9%	2.9%	4.2%	3.1%
	3+	0.7%	0.0%	0.0%	0.3%	0.4%	0.5%	0.5%	0.7%	0.4%	0.4%	0.5%	2.0%
	4+	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.2%
	5	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%

Table 4.

Comparative cumulative percent with the specified number of adjusted executive function low scores below the specified percentile cutoff (n=506)

Number of low scores		**Cumulative percent**		**Chi-Square Test**		**Wilcoxon Signed-Rank Test**		
		Latin-Americans	**U.S.A.**	X^2	***p* value**	**V**	**Z**	***p* value**
<25th percentile	None	43.6%	0.0%	329.97	<.001	328	-11.467	<.001
	1+	56.4%	100.0%					
	2+	36.0%	78.2%					
	3+	18.8%	30.1%					
	4+	6.3%	2.4%					
	5	2.2%	0.0%					
<16th percentile	None	60.8%	0.2%	482.48	<.001	388	-15.839	<.001
	1+	39.2%	99.8%					
	2+	19.0%	64.4%					
	3+	9.1%	17.2%					
	4+	4.0%	0.6%					
	5	0.4%	0.0%					
<10th percentile	None	76.0%	1.0%	598.13	<.001	412	-17.987	<.001
	1+	24.0%	99.0%					
	2+	12.7%	47.1%					
	3+	4.6%	9.1%					
	4+	0.8%	0.2%					
	5	0.0%	0.0%					
<5th percentile	None	84.4%	3.0%	671.69	<.001	427	-16.619	<.001
	1+	15.6%	97.0%					
	2+	7.5%	32.3%					
	3+	1.6%	3.6%					
	4+	0.2%	0.2%					
	5	0.0%	0.0%					
<2th percentile	None	91.9%	11.9%	635.75	<.001	414	-20.005	<.001
	1+	8.1%	88.1%					
	2+	2.6%	20.0%					
	3+	0.4%	1.2%					
	4+	0.0%	0.0%					
	5	0.0%	0.0%					

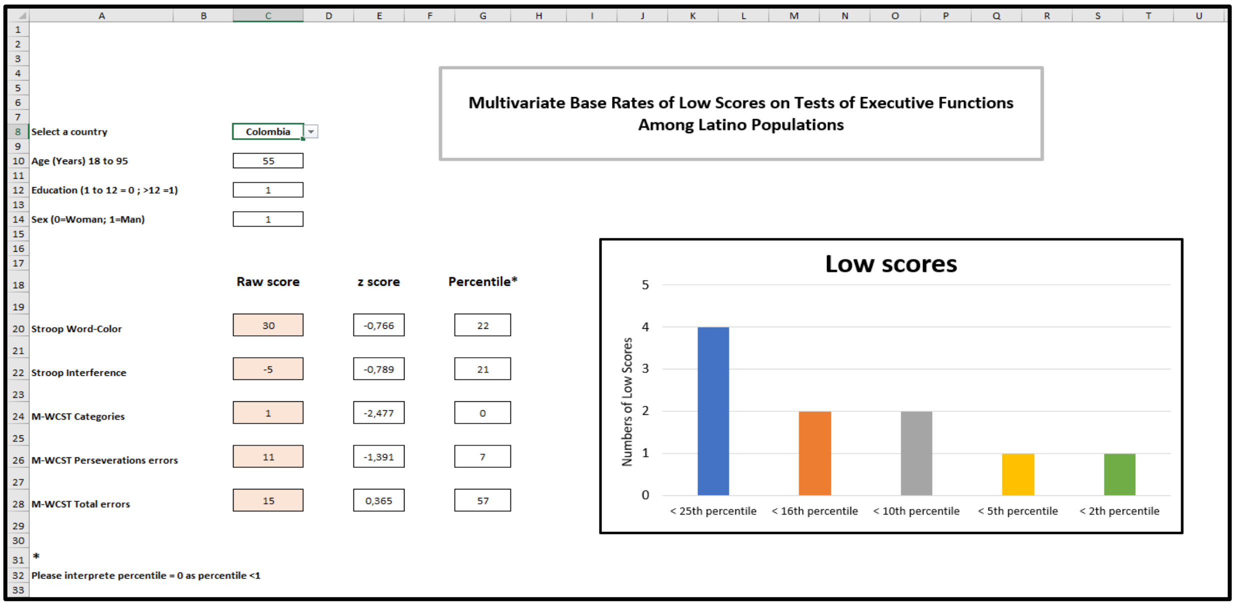

Figure 1. Calculator of low scores.