Prevalence of Low Scores on Executive Functions Tests in a Spanish-Speaking Pediatric Population from 10 Latin American Countries and Spain

Itziar Benito-Sánchez[a,b], Isabel Gonzalez[c], Rafael E. Oliveras-Rentas[d], Rosario Ferrer-Cascales[e], Ivonne Romero-García[f], Juan Carlos Restrepo Botero[g], Ivan Darío Delgado-Mejía[h], Esperanza Vergara-Moragues[i], Diego Rivera[j] and Juan Carlos Arango-Lasprilla[a,k,l*]

[a]*Biocruces Bizkaia Health Research Institute. Cruces University Hospital, Barakaldo, Spain;* [b]*Biomedical Research Doctorate Program, University of the Basque Country. Leioa, Spain;* [c]*Children's Hospital Colorado, United States;* [d]*Ponce Health Science University, Ponce, Puerto Rico;* [e]*Department of Health Psychology. University of Alicante, Spain;* [f]*Inter American University of Puerto Rico, San German Campus, San Germán, Puerto Rico;* [g]*Tecnológico de Antioquia - Institución Universitaria, Medellín, Colombia;* [h]*Universidad Autónoma de Asunción, Paraguay;* [i]*Departamento de Psicobiología y Metodología en Ciencias del Comportamiento, Universidad Complutense de Madrid (UCM), Spain;* [j]*Departamento de Ciencias de la Salud, Universidad Pública de Navarra, Pamplona, Spain;* [k]*IKERBASQUE. Basque Foundation for Science, Bilbao, Spain;* [l]*Department of Cell Biology and Histology, University of the Basque Country. Leioa, Spain*

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

Prevalence of Low Scores on Executive Functions Tests in a Spanish-Speaking Pediatric Population from 10 Latin American Countries and Spain

Pediatric neuropsychologists and researchers commonly interpret a low score as a cognitive weakness. The purpose of the study was to determine the prevalence of low scores for three neuropsychological tests used to evaluate executive function in 4,595 healthy children from Latin-America and Spain. Results showed that low scores are common when multiple neuropsychological outcomes are evaluated in healthy individuals. Clinicians should consider the higher probability of low scores in a given individual when evaluating executive functions using various sets of scores to reduce false-positive diagnoses of cognitive deficits in a child.

Key Words: Neuropsychological test, psychometrics, Diagnosis, Children, Executive Functioning test.

Introduction

Pediatric neuropsychologists and researchers commonly interpret a low score as a cognitive weakness. However, regardless of the culture and specific measure, a significant percentage of healthy children and adolescents will obtain remarkably low tests scores (Brooks, 2010). In the general population, this is a surprising finding, but in a diverse population, which by default already scores lower in most neuropsychological measures, it is of extreme importance. Recent studies in a sample of healthy individuals found that those with a lower educational level (Steinberg, Bieliauskas, Smith, & Ivnik, 2005) and intellectual ability (Brooks, Iverson, & Holdnack, 2013a) are more likely to obtain low scores in neuropsychological assessments. Interestingly, a new line of research led by Brooks and colleagues (Brooks, 2010; Brooks et al., 2013a; Brooks, Sherman, & Iverson, 2010; Brooks, Strauss, Sherman, Iverson, & Slick, 2009a), mainly using a sample of white Caucasians or Canadians, suggests that a low score in one test is not directly representative of a weakness or deficit and could instead lead to a false-positive interpretation (Brooks et al., 2013a).

Furthermore, statistical data proposes that the more tests included in an assessment, the higher the possibility of finding one or more low scores (increasing the rate of false-positive interpretations) (Brooks, 2010; Brooks et al., 2013a). For example, Brooks (2010) found that using the Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV; Wechsler, 2003) in an American sample, 1 out of 5 children and adolescents had at least one low subtest score (2 standard deviations below the mean). Brooks, Iverson, Sherman & Holdnack (2009b) also discovered in the Children's Memory Scale (CMS; Cohen, 1997), that 22% of healthy children obtained one or more index scores below the 5th percentile and when six index scores were considered 37% of children had one or more index scores below the 16th percentile. Similar results were

found for the TVCF (Reynolds & Horton, 2006), 38% of healthy children and adolescents had one or more scores below the 16th percentile and 10% had two or more low scores (Brooks et al., 2013b). For the NEPSY-Second Edition (NEPSY-II; Korkman, Kirk, & Kemp, 2007), 25% of 3 – 4-year-olds, 21% to 30% of 5 – 6-year-olds, and 35% to 44% of 7 – 16-year-olds obtained one or more scores below the 5th percentile (Brooks et al., 2010).

Brooks, Holdnack and Iverson (2017) indicated a more reliable approach is the interpretation of the base rates of low scores in a sample of healthy individuals or Multivariate Base Rates (MVBR). The use of MVBR analysis controls for false-positive interpretations, decreasing the likelihood of misdiagnosing a cognitive impairment (Brooks, 2010). This approach has been used to study several adult and pediatric neuropsychological batteries using a variety of Caucasian and Canadian samples. However, there is no data regarding normal variability in pediatric neuropsychological measures for Spanish-speaking children and adolescents.

Most Spanish speaking pediatric neuropsychologist include in their evaluations a variety of EF (Executive function) tasks. EF is described as goal-directed cognitive functions (Welsh & Pennington, 1988) and is believed to encompass organization, planning, flexibility, sequencing, fluency, switching, inhibition, concept formation, estimation, prediction, and sustained attentional sets (Banich, 2004). These higher-order cognitions are considered to be responsible for many cognitive functions related to the prefrontal cortex of the brain (Bialystok, Craik, & Luk, 2012; Iampietro, Giovannetti, Drabick, & Kessler, 2012). Deficits in EF are associated with several developmental disorders such as attention deficit/hyperactivity disorder (ADHD), Tourette syndrome, and autism (Barkley, 1998), as well as other neurological conditions such as acquired brain injuries (e.g. traumatic brain injury, epilepsy, stroke, hypoxia; Golden, 2007),

posttraumatic stress disorder (PTSD; Bremner et al., 2004), depression, and chronic pain (Golden, 2007). Due to the importance of higher order skills in everyday functioning, García-Molina, Tormos, Bernabeu, Junque & Roig-Rovira (2012) found that impairments in EF are significantly correlated with barriers for functional independence, social participation and productive living.

Given the importance of the evaluation of EF in pediatric populations and lack of knowledge of the normal variability of the low scores on Spanish Speaking children and adolescents, the aim of this study was to establish the prevalence of low scores in three EF measures in a large pediatric Spanish-speaking sample.

Method

Participants

The sample consisted of 4,595 healthy children and adolescents who were recruited from private and public schools in Chile, Colombia, Cuba, Ecuador, Guatemala, Honduras, Mexico, Paraguay, Peru, Puerto Rico, and Spain. The demographic characteristics (age, sex, type of school and mean level of parental education) by country can be found in Table 1.

Insert Table 1

To be eligible for study participation, participants must have met the following requirements: a) being between the ages of 6 and 14 years old, b) being born and currently living in the country where the protocol was administered, c) having Spanish as primary language, d) having an Intelligence Quotient (IQ) of ≥80 according to the Test of Non-Verbal Intelligence (TONI-2; Brown, Sherbenou, & Johnsen, 2009), e) having a score of

<19 on the Children's Depression Inventory (CDI, Kovacs, 1992), and f) being enrolled in a regular private or public school.

The parents (or guardians) of each potential child or adolescent participants answered a sociodemographic questionnaire about the subjects' medical history and health status. Participants were excluded according to the following criteria: a) having a history of a central nervous system disease that is associated with neuropsychological problems (e.g. epilepsy, brain injury, movement disorders, multiple sclerosis, brain tumor, stroke); b) having a history of alcohol abuse and/or consumption of psychotropic substances; c) having some type of active or uncontrolled systemic disease associated with cognitive impairment (e.g. diabetes mellitus, hypothyroidism, vitamin B12 deficiency); d) having a history of psychiatric illness (e.g. mayor depression, bipolar mood disorder, psychosis); e) having severe sensory deficits (e.g. loss of vision and/or hearing) that affect the administration of or performance on the tests; f) being on a psychiatric or other medications that could alter cognitive performance; g) having an intellectual or learning disability or other neurodevelopmental disorders; h) having a history of pre- peri-, and post-natal problems (e.g. hypoxia, jaundice, seizures, hydrocephalus, spine bifida, neuromuscular disorders); i) having a score of >5 on the Alcohol Use Disorders Identification Test (AUDIT-C) for participants 12 years of age and older; and j) using psychoactive substances such as heroin, barbiturates, amphetamines, methamphetamines, or cocaine in the last 6 months for participants 12 years of age and older. For further information regarding study's procedure, see Rivera, & Arango-Lasprilla (2017), and Arango-Lasprilla, Rivera, & Olabarrieta-Landa (2017c). The centers that agreed to participate in the study requested approval from their institution/center's ethics committee.

Measures

The Trail Making Test (TMT) consists of two parts in which the task is to connect randomly distributed points. In part A, all objectives are numbers (1 to 15) and the participant must join them up (1-2-3 …), while in part B, must alternate between numbers and letters of the next lever (1-A, 2-B, 3-C…). The total score is the number of seconds that the participant needs to complete the task (Reitan, 1992). For this analysis, only the TMT-B was used.

The Stroop Color-Word Interference 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). For this analysis Total Word-Color and Interference index were taken into account.

The Modified Wisconsin Card Sorting Test (M-WCST) consists of four stimulus cards and 48 response cards. Each card varies in shape, color, and number. The objective is to classify correctly the stimulus card according to a 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 participant responded the TMT, Stroop Word Color Test and M-WCST as part of a large battery of neuropsychological tests.

Statistical Analyses

Demographic variables effect on neuropsychological performance

The effects of demographic variables on TMT-B, Stroop (Total Word-Color and Interference) and M-WCST (number of categories, perseverations, and total errors) scores were evaluated using multiple linear regression analyses. The full regression models included age, age2, sex, mean level of parental education (MLPE), and all two-way interactions between these variables as predictors. Age was centered (= calendar age – mean age in the sample by country) before computing the quadratic age to avoid multicollinearity (Kutner, Nachtsheim, Neter, & Li, 2005). MLPE was coded variable with 1 if the participant(s) parents had > 12 mean years of education, otherwise as 0 (Guàrdia-Olmos, Peró-Cebollero, Rivera & Arango-Lasprilla, 2015; Peña-Casanova et al., 2009), and sex was coded as boys = 1 and girls = 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 as long as it was also included in a higher order term in the model (Aiken, West & Reno, 1991). The full regression model can be formally described as:

$$Eq.\ 1 \quad y_i = B_0 + B_1 \cdot (Age - \bar{x}_{Age\ by\ country})_i + B_2 \cdot \left(Age - \bar{x}_{Age\ by\ country}\right)_i^2 + B_3 \cdot Sex_i + B_4 \cdot MLPE_i.$$

Calculation of adjusted Z-score

The effects of demographic variables previously described on the scores were adjusted and converted to z-scores using the final regression model published by Arango-Lasprilla et al. (2017c) for a Colombian sample, Arango-Lasprilla et al. (2017b; TMT-B scores), Rivera et al. (2017; Stroop scores) and Arango-Lasprilla et al. (2017a; M-WCST scores)

for the other countries (Chile, Cuba, Ecuador, Guatemala, Honduras, Mexico, Paraguay, Peru, Puerto Rico and Spain).

This 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): 1. The expected test score ($\hat{Y}_i$) is computed based on the fixed effect parameter estimated for the established final regression model in Eq. 1; 2. To obtain the residual value (e_i), predicted values ($\hat{Y}_i$) are subtracted from raw scores (Y_i) as shown in the following formula: $e_i = Y_i - \hat{Y}_i$. 3. Residuals are standardized using the residual standard deviation (SD_e) of the regression model: $z_i = e_i / SD_e$. This three-step process was applied to each score (TMT-B, Stroop Total Word-Color, Stroop Interference, M-WCST number of categories, M-WCST perseverations and M-WCST total errors) separately for each country.

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.

For each subject, the number of tests at or below each of the percentiles was calculated. The cumulative proportion of k or more low tests was used as the base rate of low scores and involved examination of performance on the six measures (TMT-B, Stroop Total Word-Color, Stroop Interference, M-WCST number of categories, M-

WCST perseverations and M-WCST total errors) simultaneously, not each score in isolation. All analyses were performed using SPSS Version 23.0 (IBM Corp., 2015). The number of k or more low tests found in fewer than 10% of the normative sample was used as a criterion for an unusual number of low tests, as used in past research (Binder et al., 2009; Mistridis et al., 2015; Oltra-Cucarella et al., 2018).

Results

The base rates of low scores on the executive functioning performance after adjusting for age, gender and MLPE, are presented in Table 2.

Between 55.9% (Chile) and 70.2% (Guatemala) of the sample have at least 1 score below the 25th percentile and between 42.5% (Paraguay) and 52.0% (Guatemala) scored below the 16th percentile. Moreover, between 32.4% (Paraguay) and 41.4% (Honduras) scored below the 10th percentile on at least 1 of the 6 sub-tests and between 21.1% (Guatemala) and 27.3% (Ecuador) scored below the 5th percentile. Finally, between 11.6% (Mexico) and 18.3% (Puerto Rico) scored below the 2nd percentile on at least 1 of the 6 scores.

Insert Table 2

Examples will be provided to facilitate the interpretation of Table 2. For example, in Paraguay, 62.6% of the sample have at least 1 or more scores below the 25th percentile, 42.5% below the 16th percentile, 32.4% below the 10th percentile, 21.2% below the 5th percentile, and 12.3% below the 2nd percentile. The results are represented in Figure 1.

Insert Figure 1

In another example, in Honduras, 61.7% of the sample have at least 1 or more scores below the 25th percentile, 51.1% below the 16th percentile, 41.4% below the 10th percentile, 25.1% below the 5th percentile, and 13.7% below the 2nd percentile. The results are represented in Figure 2.

Insert Figure 2

The number of low scores that contain less than 10% of the normative sample and are considered to be unusual numbers of low scores is consistent across countries for each percentile. The unusual numbers of low scores are four or more for the 25th and 16th percentiles (three or more for Cuba, Paraguay and Peru), three or more for the 10th percentiles (except for Ecuador), between three and two or more for the 5th percentile, and two or more for the 2nd percentile. See Table 2.

Discussion

Until recently, the availability of tests to evaluate EF in children and adolescents with country-specific normative data were limited for most Spanish-speaking countries. After a multi-country study, normative data for 10 commonly used neuropsychological tests in children and adolescents are now available for Chile, Colombia, Cuba, Ecuador, Guatemala, Honduras, Mexico, Paraguay, Peru, Puerto Rico, and Spain (Arango-Lasprilla et al., 2017c; Rivera, & Arango-Lasprilla, 2017). Therefore, one of the major problems reported by many practitioners in providing neuropsychological services for Spanish Speakers (Arango-Lasprilla, Stevens, Morlett Paredes, Ardila, & Rivera, 2017d)

was addressed. Moreover, using the present base rate analyses would provide even more objective information in interpreting test performance, reducing the likelihood of false positives misinterpretation (i.e. interpreting a low score as impairment to a cognitively unimpaired individual). To our knowledge, this is the first cross-country analysis of base rates in test scores for pediatric Spanish Speakers.

The results of the present study are generally consistent with previous research, in terms that low scores were very common among the sample. In this study, when stratified by country, and after controlling by age, gender and mean level of parental education, a minimum of 55.9% (Chile) of the individuals have at least one score below the 25th percentile of the six measures of EF analyzed. Even a range from 42.5% (Paraguay) to 52.0% (Guatemala) of subjects presented one score below the <16th percentile. From a statistical perspective, these numbers suggest that approximately half of the children would present one score below a standard deviation from the mean when these six EF measures are considered (i.e. Stroop, TMT-B and M-WCST). This challenges the interpretation from the statistical assumption of normality for a single score, which establishes that 68% of the sample would score in a test within a standard deviation from the mean. In order to consider the 16th percentile as a cut-off for impairment, an interpretation of three or more scores would be required, as found in 6.1-15.2% across the countries. If a 5th percentile is used as a cut-off score to consider impairment, it should be noted that a range between 21.1% (Guatemala) and 27.3% (Ecuador) of healthy children across countries have at least one or more score at or below this level. Findings of a high prevalence of low scores when interpreting multiple tests scores have been reported in the literature on several test batteries for children, including the Children's Memory Scale (Brooks et al., 2009a), NEPSY-II (Brooks et al., 2010) and with different samples on the Wechsler Intelligence Scale for Children, Fourth Edition (Brooks, 2010).

In addition, Brooks et al. (2013b) also found an elevated percentage of low scores when using five executive function measures on the Test of Verbal Conceptualization and Fluency (TVCF). On that sample, composed by 719 children and adolescents between 8 and 19 years old, 38% had one or more low scores using the 16th percentile cut-off score, and 15% had at least one low score at or below the 5th percentile (Brooks et al., 2013b).

Also, consistent with previous research, age had limited impact on the base rates. In this study, there was not a specific pattern in terms of percentage of children and adolescents obtaining one or more low scores at different percentile cut-off scores when age range was divided between 6 to 8, 9 to 11, and 12 to 14 years. Brooks and colleagues (2009b) found the same results while evaluating the frequency of low scores across memory tests. In fact, their findings were generally consistent with literature using adult memory tests, suggesting that the presence of low scores when interpreting multiple scores, is a psychometric principle independent of cognitive development (Brooks et al., 2009b). Therefore, low scores prevalence analyses can be considered a validated practice for test interpretation in developmental neuropsychology.

Parental education level, however, has been reported as influential factor in the prevalence of low scores in children. Brooks et al. (2010) found fewer rates of low scores on the NEPSY-II on children when their parents had more years of education, and vice versa. This relatively new approach of stratifying base rates by parental education emerges from the evidence of the strong correlation between parental education and cognitive function in children (Schoenberg, Lange, & Saklofske, 2007; Thomas, Sukumaran, Lukose, George, & Sarma, 2007; van der Sluis, Willemsen, de Geus, Boomsma, & Posthuma, 2008). In this study, most of the sample presented this tendency of a lower prevalence of low scores when MLPE >12. These findings reveal the caution that needs to be exercised when parental education is used as an independent indicator

for the estimation of the child's cognitive ability. Brooks and colleagues (2009a) described that one of the predictive factors of prevalence of low scores is general cognitive functioning, being negatively correlated. Thus, if cognitive functioning is influenced in part by parental education, then MLPE can be considered as an indirect factor for low score prevalence. Brooks (2010) and Brooks et al. (2010) recommended that both general functioning and parental education should be considered before drawing conclusions in terms of a child's functioning compared to healthy children. In this study, intelligence was measured with the TONI-2, but only as part of the inclusion criteria. Another explanation of those exceptions can be related to the nature of the EF measured. Studies suggest that there are EF subcomponents that does not seem to correlate with intelligence, including inhibition and attention shifting (Friedman et al., 2006; Was, 2008), which are presumably measured by Stroop, TMT and M-WCST. This area clearly needs further investigation.

There are some limitations with this study. First, the sample was not stratified by intellectual ability, as it was used only as an inclusion criterion for this study. To further understand the role of intelligence, MLPE and EF measures, future research should explore the relationship between those factors within this pediatric sample. In addition, although each family answered a detailed sociodemographic questionnaire, the sample was not formally screened for EF deficits or specific disorders.

Another consideration is that these base rates of low scores should be interpreted if only those 6 measures of executive functioning are used. Scores gathered from the sample were part of a more comprehensive battery of selected tests. It is estimated that if more scores are used, the prevalence of low scores would be higher (Brooks et al., 2010).

In addition, the present study was conducted with children and adolescents from 11 Spanish Speaking countries, in which different base rate percentages were derived

from each country, age and MLPE. Because of the variability seen, it is likely that these results cannot be generalized to other countries.

This study demonstrated that low scores are a common feature in healthy individuals -even after adjusted by sociodemographic variables- when several measures are administered. Pediatric neuropsychologists and related practitioners who work with Spanish Speaking children and adolescents now count with the base rates of low scores from six common EF measures from 11 Latino America and Spain countries. This psychometric tool should encourage practitioners to avoid interpreting single test scores in isolation, as it raises the likelihood of a false positive interpretation. Future research should examine the prevalence of low scores in various clinical samples, particularly considering the nature, development and subcomponents of EF.

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., & Olabarrieta-Landa, L. (Eds.). (2017c). *Neuropsicología infantil*. Manual Moderno.

Arango-Lasprilla, J. C., Rivera, D., Nicholls, E., Aguayo Arelis, A., García de la Cadena, C., Peñalver Guia, A. I., ... & Sánchez-SanSegundo, M. (2017a). Modified Wisconsin Card Sorting Test (M-WCST): Normative data for Spanish-speaking pediatric population. *NeuroRehabilitation, 41*(3), 617-626.

Arango-Lasprilla, J. C., Rivera, D., Ramos-Usuga, D., Vergara-Moragues, E., Montero-López, E., Adana Díaz, L. A., … & Ferrer-Cascales, R. (2017b). Trail Making Test: Normative data for the Latin American Spanish-speaking pediatric population. *NeuroRehabilitation, 41*(3), 627-637.

Arango-Lasprilla, J. C., Stevens, L., Morlett Paredes, A., Ardila, A., & Rivera, D. (2017d). Profession of neuropsychology in Latin America. *Applied Neuropsychology: Adult, 24*(4), 318-330.

Banich, M.T. (2004). *Cognitive neuroscience and neuropsychology.* Second Edition. Boston (Massachusetts): Houghton Mifflin.

Barkley, R. A. (1998). Symptoms, diagnosis, prevalence, and gender differences. In R. A. Barkley (Ed.) *Attention deficit/hyperactivity disorder: A handbook for diagnosis and treatment* (2nd ed.). New York: Guilford Press.

Bialystok, E., Craik, F. I., & Luk, G. (2012). Review: Bilingualism: Consequences for mind and brain. *Trends In Cognitive Sciences, 16*(4), 240-250. doi:10.1016/j.tics.2012.03.001.

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. doi:10.1093/arclin/acn001

Bremner, J. D., Vermetten, E., Vythilingam, M., Afzal, N., Schmahl, C., Elzinga, B., & Charney, D. S. (2004). Neural correlates of the classic color and emotional stroop in women with abuse-related posttraumatic stress disorder. *Biological psychiatry, 55*(6), 612-620.

Brooks, B. L. (2010). Seeing the forest for the trees: Prevalence of low scores on the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV). *Psychological Assessment*, *22*(3), 650–656.

Brooks, B. L., Holdnack, J. A, & Iverson, G. L. (2017). Reliable Change on Memory Tests is Common in Healthy Children and Adolescents, *Archives of Clinical Neuropsychology*, *32*(8), 1001-1009. doi:10.1093/arclin/acx028.

Brooks, B. L., Iverson, G. L., & Holdnack, J. A. (2013a). Understanding multivariate base rates. In Holdnack J. A., Drozdick L., Weiss L. G., & Iverson G. L. (Eds.), WAIS-IV, WMS-IV, & ACS: *Clinical use and interpretation* (pp. 75–102). New York: Elsevier.

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

Brooks, B. L., Iverson, G. L., Sherman, E. M. S., & Holdnack, J. A. (2009b). Healthy children and adolescents obtain some low scores across a battery of memory tests. *Journal of the International Neuropsychological Society*, *15*(4), 613–617.

Brooks, B. L., Sherman, E. M., & Iverson, G. L. (2010). Healthy children get low scores too: Prevalence of low scores on the NEPSY-II in preschoolers, children, and adolescents. *Archives of Clinical Neuropsychology, 25*(3), 182-190. https://doi.org/10.1093/arclin/acq005.

Brooks, B. L., Strauss, E., Sherman, E. M. S., Iverson, G. L., & Slick, D. J. (2009a) Developments in neuropsychological assessment: Refining psychometric and clinical interpretive methods. *Canadian Psychology/Psychologie canadienne, 50*(3), 196-209.

Brown L., Sherbenou R. J., & Johnsen S. K. (2009), *Test de inteligencia no verbal TONI-2*. Madrid: TEA ediciones.

Cohen, M.J. (1997). *Children's Memory Scale (CMS) Manual*. The Psychological Corporation, San Antonio.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., Defries, J. C., & Hewitt, J. K. (2006). Not All Executive Functions Are Related to Intelligence. *Psychological Science, 17*(2), 172-179.

García-Molina, A., Tormos, J., Bernabeu, M., Junque, C., & Roig-Rovira, T. (2012). Do traditional executive measures tell us anything about daily-life functioning after traumatic brain injury in Spanish-speaking individuals? *Brain Injury*, *26*(6), 864-874. doi:10.3109/02699052.2012.655362.

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

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

Iampietro, M., Giovannetti, T., Drabick, D., & Kessler, R. (2012). Empirically defined patterns of executive function deficits in schizophrenia and their relation to everyday functioning: A person-centered approach. *The Clinical Neuropsychologist*, *26*(7), 1166-1185. doi:10.1080/13854046.2012.721399.

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

Korkman, M., Kirk, U., & Kemp, S. (2007). *NEPSY-II: A developmental neuropsychological assessment*. San Antonio, TX: The Psychological Corporation.

Kovacs, M. (1992). *Children's depression inventory*. North Tonawanda, NY: Multi-Health System.

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

Mistridis, P., Egli, S. C., Iverson, G. L., Berres, M., Willmes, K., Welsh-Bohmer, K. A., & Monsch, A. U. (2015). Considering the base rates of low performance in cognitively healthy older adults improves the accuracy to identify neurocognitive impairment with the Consortium to Establish a Registry for Alzheimer's Disease-Neuropsychological Assessment Battery (CERAD-NAB). *European Archives of Psychiatry and Clinical Neuroscience, 265*(5), 407-417. doi:10.1007/s00406-014-0571-z

Oltra-Cucarella, J., Sánchez-SanSegundo, M., Lipnicki, D. M., Sachdev, P. S., Crawford, J. D., Pérez-Vicente, J. A., … & Ferrer-Cascales, R. (2018). Using the base rate of low scores helps to identify progression from amnestic MCI to AD. *Journal of the American Geriatrics Society*, *66*(7), 1360–1366. doi:10.1111/jgs.15412

Peña-Casanova, J., Blesa, R., Aguilar, M., Gramunt-Fombuena, N., Gómez-Ansón, B., Oliva, R., … & Martínez-Parra, C. (2009). Spanish multicenter normative studies

(NEURONORMA project): Methods and sample characteristics. *Archives of Clinical Neuropsychology*, *24*(4), 307-319.

Reitan R. M. (1992), *Trail Making Test: Manual for Administration and Scoring*. Tucson, AZ: Reitan Neuropsychology Laboratory

Reynolds, C., & Horton, A. M., Jr. (2006). *Test of Verbal Conceptualization and Fluency*. Austin, TX: Pro-Ed.

Rivera, D., & Arango-Lasprilla, J. C. (2017). Methodology for the development of normative data for Spanish-speaking pediatric populations. *NeuroRehabilitation, 41*(3), 581-592.

Rivera, D., Morlett-Paredes, A., Peñalver Guia, A. I., Irías Escher, M. J., Soto-Añari, M., Aguayo Arelis, A., ... & García de la Cadena, C. (2017). Stroop Color-Word Interference Test: Normative data for Spanish-speaking pediatric population. *NeuroRehabilitation, 41*(3), 605-616.

Schoenberg, M. R., Lange, R. T., & Saklofske, D. H. (2007). A proposed method to estimate premorbid full scale intelligence quotient (FSIQ) for the Canadian Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV) using demographic and combined estimation procedures. *Journal of Clinical and Experimental Neuropsychology, 29*(8), 867–878.

Schretlen D. (2010). *Modified Wisconsin Card Sorting Test: M-WCST*. Professional Manual. PAR.

Steinberg, B. A., Bieliauskas, L. A., Smith, G. E., & Ivnik, R. J. (2005). Mayo's older Americans normative studies: Age- and IQ-Adjusted Norms for the Trail-Making Test, the Stroop Test, and MAE Controlled Oral Word Association Test. *The Clinical Neuropsychologist*, *19*(3-4), 329–377.

Thomas, S. V., Sukumaran, S., Lukose, N., George, A., & Sarma, P.S. (2007).

Intellectual and language functions in children of mothers with epilepsy.

Epilepsia, 48(12), 2234–2240.

van der Sluis, S., Willemsen, G., de Geus, E. J., Boomsma, D. I., & Posthuma, D.

(2008). Gene-environment interaction in adults' IQ scores: Measures of past and

present environment. *Behavioral Genetics, 38*(4)*, 348–360.

Was, C. A. (2008). Further evidence that not all executive functions are equal. *Advances

in cognitive psychology*, *3*(3), 399-407. doi:10.2478/v10053-008-0004-5

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children-Fourth Edition.* San

Antonio, TX: The Psychological Corporation.

Welsh, M. C., & Pennington, B. F. (1988). Assessing frontal lobe functioning in children:

Views from developmental psychology. *Developmental Neuropsychology, 4*(3),

199-230.

Table 1. Sample distribution by country, age, sex, type of school, and MLEP.

	N	Age Mean (SD)	Sex Girls n (%)	Boys n (%)	Type of school Public n (%)	Private n (%)	MLPE Mean (SD)
Chile	289	9.9 (2.5)	145 (50.2%)	144 (49.8%)	145 (50.2%)	144 (49.8%)	12.5 (2.9)
Colombia	1262	10.1 (2.5)	661 (52.4%)	601 (47.6%)	636 (50.4%)	626(49.6%)	12.3 (3.6)
Cuba	286	10.0 (2.5)	142 (49.7%)	144 (50.3%)	286 (100%)	0 (0.0%)	16.1 (1.8)
Ecuador	230	9.9 (2.5)	134 (58.3%)	96 (41.7%)	119 (51.7%)	111 (48.3%)	14.4 (3.4)
Guatemala	181	10.1 (1.9)	79 (43.9%)	101 (56.1%)	111 (61.3%)	70 (38.7%)	10.0 (3.9)
Honduras	240	10.0 (2.4)	127 (52.9%)	113 (47.1%)	137 (57.1%)	103 (42.9%)	12.6 (3.7)
Mexico	707	9.8 (2.6)	362 (51.2%)	345 (48.8%)	420 (59.4%)	287 (40.6%)	13.4 (3.9)
Paraguay	214	9.9 (2.5)	120 (56.1%)	94 (43.9%)	103 (48.1%)	111 (51.9%)	14.1 (2.8)
Peru	254	10.4 (2.5)	118 (46.5%)	136 (53.5%)	125 (49.2%)	129 (53.5%)	12.4 (2.3)
Puerto Rico	148	10.4 (2.8)	84 (56.8%)	64 (43.2%)	88 (59.5%)	60 (40.5%)	14.6 (2.4)
Spain	784	9.9 (2.5)	397 (50.6%)	387 (49.4%)	428 (54.6%)	356 (45.4%)*	14.1 (4.0)

Note: MLPE: Mean Level Parental Education; *Private/Concerted (private school

partially publicly funded).

Table 2. Cumulative proportion of children with the specified number of adjusted executive functioning low scores below the specified percentile cutoff by country.

Executive Function		All countries	Chile	Colombia	Cuba	Ecuador	Guatemala	Honduras	Mexico	Paraguay	Peru	Puerto Rico	Spain
<25th percentile	None	40.0%	44.1%	42.1%	35.7%	40.9%	29.8%	38.3%	40.3%	37.4%	38.1%	39.4%	40.7%
	1+	60.0%	55.9%	57.9%	64.3%	59.1%	70.2%	61.7%	59.7%	62.6%	61.9%	60.6%	59.3%
	2+	39.0%	36.0%	37.8%	40.7%	39.4%	46.8%	43.6%	41.4%	33.5%	40.8%	29.8%	38.0%
	3+	20.6%	22.7%	19.5%	18.6%	22.7%	25.7%	20.7%	22.8%	12.3%	20.2%	17.3%	21.6%
	4+	8.2%	5.7%	8.6%	5.4%	9.8%	12.9%	9.7%	9.7%	3.9%	7.8%	4.8%	8.0%
	5+	2.7%	3.2%	2.5%	1.4%	1.5%	2.9%	3.1%	4.0%	0.6%	2.3%	--	3.1%
	6	0.4%	0.4%	0.4%	0.4%	0.8%	--	0.9%	0.6%	--	0.5%	--	0.4%
<16th percentile	None	53.3%	55.5%	54.4%	51.1%	54.5%	48.0%	48.9%	52.4%	57.5%	50.0%	51.9%	55.2%
	1+	46.7%	44.5%	45.6%	48.9%	45.5%	52.0%	51.1%	47.6%	42.5%	50.0%	48.1%	44.8%
	2+	26.0%	25.1%	25.6%	25.0%	26.5%	26.9%	29.5%	27.2%	18.4%	27.1%	22.1%	27.1%
	3+	11.3%	11.3%	10.9%	8.9%	15.2%	12.9%	10.1%	12.1%	6.1%	9.2%	10.6%	13.2%
	4+	3.4%	3.2%	3.4%	2.5%	5.3%	3.5%	3.5%	4.3%	1.7%	2.8%	1.9%	3.5%
	5+	0.9%	--	0.6%	0.7%	--	0.6%	1.8%	1.8%	--	1.4%	--	1.1%
	6	--	--	0.1%	--	--	--	--	0.2%	--	--	--	--
<10th percentile	None	64.4%	67.2%	65.1%	60.4%	63.6%	66.7%	58.6%	63.1%	67.6%	62.8%	66.3%	65.9%
	1+	35.6%	32.8%	34.9%	39.6%	36.4%	33.3%	41.4%	36.9%	32.4%	37.2%	33.7%	34.1%
	2+	16.3%	17.0%	15.8%	13.6%	19.7%	17.0%	14.1%	15.9%	10.1%	17.9%	16.3%	19.3%
	3+	6.7%	8.9%	6.0%	6.1%	11.4%	5.3%	5.3%	5.9%	5.0%	6.4%	8.7%	8.0%
	4+	1.7%	0.8%	1.5%	1.8%	1.5%	0.6%	2.2%	2.5%	1.7%	--	1.9%	1.9%
	5+	0.2%	--	0.1%	0.4%	--	--	0.4%	1.0%	--	--	--	0.1%
	6	--	--	--	--	--	--	--	--	--	--	--	--
<5th percentile	None	76.6%	75.7%	78.1%	73.2%	72.7%	78.9%	74.9%	78.0%	78.8%	72.9%	76.9%	75.7%

	1+	23.4%	24.3%	21.9%	26.8%	27.3%	21.1%	25.1%	22.0%	21.2%	27.1%	23.1%	24.3%
	2+	9.0%	10.5%	7.6%	9.3%	10.6%	7.0%	7.5%	8.1%	5.6%	11.9%	12.5%	11.3%
	3+	3.1%	4.0%	2.6%	2.9%	6.8%	0.6%	3.1%	3.2%	3.4%	2.8%	6.7%	3.1%
	4+	0.6%	0.4%	0.4%	1.1%	0.8%	--	0.4%	0.6%	1.1%	--	1.9%	0.5%
	5+	--	--	--	--	--	--	0.4%	--	--	--	--	--
	6	--	--	--	--	--	--	--	--	--	--	--	--
<2th percentile	None	86.4%	83.8%	87.3%	85.0%	85.6%	87.7%	86.3%	88.4%	87.7%	83.5%	81.7%	85.6%
	1+	13.6%	16.2%	12.7%	15.0%	14.4%	12.3%	13.7%	11.6%	12.3%	16.5%	18.3%	14.4%
	2+	3.8%	7.3%	3.1%	5.7%	3.8%	1.2%	2.6%	3.2%	3.9%	4.1%	8.7%	3.8%
	3+	1.0%	0.8%	0.8%	1.4%	1.5%	--	1.3%	0.3%	2.8%	0.9%	4.8%	0.9%
	4+	0.1%	0.4%	0.1%	1.1%	--	--	--	--	--	--	--	0.1%
	5+	--	--	--	--	--	--	--	--	--	--	--	--
	6	--	--	--	--	--	--	--	--	--	--	--	--

Figure 1. Cumulative proportion of Paraguayan children with the specified number of adjusted executive functioning low scores below the specified percentile cutoff.
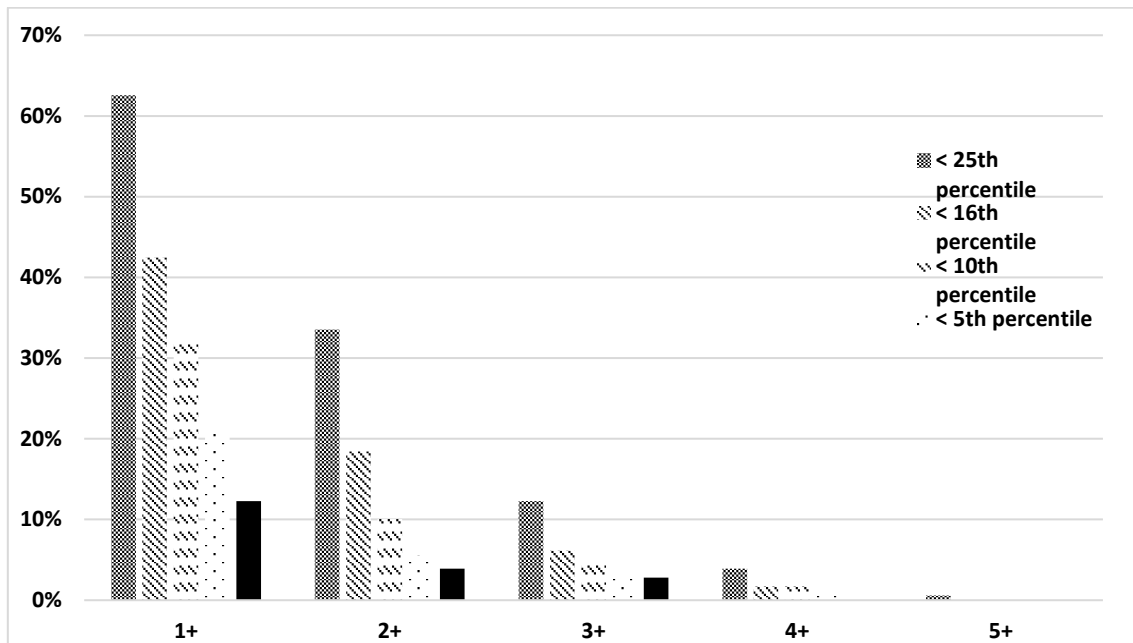
Figure 2. Cumulative proportion of Honduran children with the specified number of adjusted executive functioning low scores below the specified percentile cutoff.