Multivariate Base Rates of Low Scores on Tests of Learning and Memory Among Spanish speaking children

Itziar Benito-Sánchez[a,b], Melissa M. Ertl[c], Rosario Ferrer-Cascales[d], Javier Oltra-Cucarella[d], Joaquín A. Ibáñez-Alfonso[e], Mahia Saracostti Schwartzman[f], Marcio Soto-Añari[g], Natalia Cadavid-Ruiz[h], Alberto Rodríguez Lorenzana[i], Diego Rivera[a] 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*University at Albany, State University of New York, New York, USA; *d*Department of Health Psychology, University of Alicante, Alicante, Spain; *e*Universidad Loyola Andalucía, Department of Psychology, Sevilla, Spain; *f*Universidad de la Frontera, Temuco, Chile; *g*Departamento de Psicología, Universidad Católica San Pablo, Arequipa, Perú; *h*Pontificia Universidad Javeriana de Cali, Colombia; *i*Escuela de Psicología, Universidad de Las Américas, Quito, Ecuador; *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

Multivariate Base Rates of Low Scores on Tests of Learning and Memory Among Spanish speaking children

To determine the prevalence of low scores on two neuropsychological tests commonly used to evaluate learning and memory in children. 6,030 healthy children from 10 countries in Latin America and Spain were administered Rey–Osterrieth Complex Figure (ROCF) and the Test de Aprendizaje y Memoria Verbal–Infantil (TAMV-I). Results showed that low scores are common when multiple neuropsychological outcomes (tests and/or scores) are evaluated in healthy individuals. Clinicians should consider the higher probability of low scores in a given individual when evaluating learning and memory using various sets of scores to reduce false-positive diagnoses of cognitive deficits in pediatric populations.

Keywords: Neuropsychology, Memory and Learning test, psychometrics, diagnosis, children.

Introduction

Learning and memory, which reflect individuals' capacities for registering, consolidating, storing, and retrieving new information, are essential, complex cognitive functions. Neuropsychological assessments of learning and memory aim to measure abilities and potential deficits in these areas. When used among children and youth, neuropsychological assessments of learning and memory can discern developmental delays and relative strengths and weaknesses. Memory impairments can be common in pediatric populations with neurological diseases (Menlove & Reilly, 2015; Rayner, Jackson, & Wilson, 2016; Serra-Grabulosa, 2005), and as such, it is important to assess both memory and learning among children. Both the Rey-Osterrieth Complex Figure Test (ROCF; Rey, 1941, 2009) and the Test de Aprendizaje y Memoria Verbal Infantil (TAMV-I; Rivera, Olabarrieta-Landa, & Arango-Lasprilla, 2017a) are two neuropsychological assessment instruments that are commonly used to measure learning and memory among children and adolescents.

The ROCF measures cognitive performance by assessing recall of visual information (Fastenau, 1996). Administration of the ROCF entails presenting an asymmetrical drawn stimulus to the individual being tested. Two commonly used conditions of the ROCF include immediate and delayed recall trials (Shin, Park, Park, Seol, & Kwon, 2006). Use of the ROCF for assessing visual learning and memory among various populations, including Spanish-speaking youth, has been supported in past research (Arango-Lasprilla et al., 2017a; Waber & Holmes, 1985). Although the ROCF was originally developed for and normed with English-speaking samples, researchers later developed normative data for the ROCF for Spanish-speaking pediatric populations in Chile, Colombia, Cuba, Ecuador, Guatemala, Honduras, Mexico, Paraguay, Peru, Puerto Rico, and Spain (Arango-Lasprilla et al., 2017a; Folleco et al., 2017).

The TAMV-I was developed to assess learning and memory among Spanish-speaking samples between 6 and 17 years of age (Rivera et al., 2017a). The TAMV-I is comprised of a list of 12 words in three semantic categories. It uses three conditions: total, delayed recall, and recognition. Normative data exists for use of the TAMV-I among Spanish-speaking pediatric samples from 10 countries and the commonwealth of Puerto Rico (i.e. Chile, Colombia, Cuba, Ecuador, Guatemala, Honduras, Mexico, Paraguay, Peru, Puerto Rico, and Spain; Rivera et al., 2017b; Rivera et al., 2017c).

Standardized normative data for the population being assessed reduces the risk of misinterpretation of scores (Rivera, Olabarrieta-Landa, & Arango-Lasprilla, 2017d). In addition to normative data, another way to reduce misinterpretation of scores on neuropsychological assessments is to consult derived multivariate base rates (MVBRs) for the population. MVBRs consist of the frequency and determinants of low scores among healthy individuals in the population.

Neuropsychologists use MVBRs of low scores when interpreting large amounts of data. Chances dramatically increase for individuals to have one or more low scores on an individual test when a battery of assessments is conducted (Binder, Iverson, & Brooks, 2009; Brooks, Sherman, & Iverson, 2010) since the number of low scores is related to the number of assessments completed. When an individual receives a low score on any individual test, the clinician interpreting the results of the battery must decide whether the results reflect cognitive impairment (true positive) or a low score in an otherwise healthy individual (false positive). Other factors, including age, education, and intellectual level, may alter MVBRs and increase the prevalence of low scores (Brooks & Iverson, 2010; Schretlen, Testa, Winicki, Pearlson, & Gordon, 2008). MVBRs are a tool that can be used to improve accuracy of clinicians identifying cognitive impairments so as to reduce misdiagnosing deficits. Although MVBRs exist for preschoolers, children,

adolescents (e.g., Brooks et al., 2010), and adult English speaking populations (Schretlen et al., 2008), to date MVBRs have not been developed for use among Spanish-speaking pediatric populations.

As such, a significant gap remains in the literature regarding MVBRs among Spanish-speaking youth. To fill this gap, the present study examined MVBRs among a Spanish-speaking sample of children and adolescents from 10 countries (Chile, Colombia, Cuba, Ecuador, Guatemala, Honduras, Mexico, Paraguay, Peru, and Spain) and the commonwealth of Puerto Rico using the ROCF (immediate and delayed recall) and the TAMV-I (total, delayed recall, and recognition). Findings will be used to develop a table of the base rates of low scores on the ROCF and the TAMV-I to facilitate interpretation of scores with a focus on maintaining an adequate false-positive rate when these two assessments are administered together in a battery (e.g., Brooks et al., 2010).

It was hypothesized that the prevalence of low scores on the ROCF and TAMV-I, as determined using MVBRs, will exceed the expected prevalence rates found when interpreting a single score in isolation. Results of the present study will elucidate MVBRs among Spanish-speaking Latino children and adolescents and will improve clinical interpretation for neurological performance. Thus, findings have the distinct potential to reduce the likelihood of over-diagnosing cognitive deficits in healthy individuals.

Method

Participants

The sample included 6,030 healthy children and adolescents recruited 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 had to meet the following requirements: a) being between 6 and 17 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 $\geq$ 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 participant answered a sociodemographic questionnaire about the participant's medical history and health status. Participants were excluded according to the following criteria: a) having a history of a central nervous system disease associated with neuropsychological impairments (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 psychiatric or other medications that could alter cognitive performance; g) having 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; Bush, Kivlahan, McDonell, Fihn, & Bradley, 1998) for participants 12 years of age and older; and j) Self-reported history of use of psychoactive substances, such as heroin, barbiturates, amphetamines, methamphetamines, or cocaine in the last 6 months for participants 12 years of age and older. The centers that agreed to participate in the study requested approval from their institution/center's ethics committee. For further information regarding study's procedure, see Rivera and Arango-Lasprilla (2017) and Rivera et al.(2017d).

Measures

Rey–Osterrieth Complex Figure (ROCF). The examiner administered the ROCF Figure A (copy and 3-minutes delayed recall). The Spanish-language ROCF manual scoring guidelines were followed (Rey, 2009). The ROCF includes 18 elements, with a score range for the two tasks between 0 and 36. Two points are given when the element is correctly reproduced; 1 point when the reproduction is distorted, incomplete but placed properly, or complete but placed poorly; and 0.5 point is credited when the element is distorted or incomplete and placed poorly. A score of 0 is given when the element is absent or is not recognizable (Osterrieth, 1944).

Test de Aprendizaje y Memoria Verbal Infantil (TAMV-I). The TAMV-I consists of a list of 12 words belonging to three semantic categories: clothing, furniture, and body parts. The test is administered in 4 trials. In each trial, first the examiner tells the examinee to repeat the 12 words that will be read aloud. Immediately after reading the 12 words, the examinee is asked to recall as many words as possible. All 4 trials of listing and remembering must be completed. Therefore, there is an immediate recall total learning score with a maximum of 48 points (12 per trial). After 30 minutes, the examinee is asked to remember and list the same 12 words. Thus, the total memory delayed recall score has a maximum of 12 points. After the delayed recall trial, the recognition trials begin. In

each of 12 recognition trials, the examinee must identify the initial words from the list of 12 when presented with 3 other words in groups of 4: 1 semantically related to the initial word, 1 non-semantically related, and 1 phonologically related. Thus, the total recognition score has a maximum of 12 points. The TAMV-I administration guidelines can be found in Rivera et al. (2017a). The participant performed the ROCF and TAMV-I as part of a large battery of neuropsychological tests (see Rivera & Arango-Lasprilla, 2017).

Statistical Analyses

Demographic variables effect on neuropsychological performance

The effects of demographic variables on ROCF (immediate and delayed recall) and TAMV-I (total recall, delayed recall, and recognition) scores were evaluated using multiple linear regression analyses. The full regression models included age, age^2, 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 models published by Folleco et al. (2017; ROCF scores) and Rivera et al. (2017b; TAMV-I scores) for Colombian sample, and by Arango-Lasprilla et al. (2017a; ROCF scores) and Rivera et al. (2017c; TAMV-I scores) for the other countries (Chile, Cuba, Ecuador, Guatemala, Honduras, Mexico, Paraguay, Peru, Puerto Rico and Spain).

These adjusted z-scores for each raw score were calculated 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 of 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 (ROCF immediate recall, ROCF delayed recall, TAMV-I total, TAMV-I delayed recall, and TAMV-I recognition) 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 five measures (ROCF immediate recall, ROCF delayed recall, TAMV-I total, TAMV-I delayed recall, and TAMV-I recognition) 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 learning and memory performance after adjusting for age, gender and MLPE, are presented in Table 2.

Between 50.1% (Peru) and 61.9 % (Chile) of the sample have at least 1 of the 5 scores below the 25th percentile, and between 37.4% (Peru) and 48.9% (Puerto Rico) below the 16th percentile. Moreover, 27.3% (Peru) and 39.2% (Puerto Rico) scored below the 10th percentile on at least 1 of the 5 sub-tests, and between 19.9% (Peru) and 27.3% (Puerto Rico) below the 5th percentile. Finally, between 10.7% (Honduras) and 16.1% (Cuba) scored below the 2nd percentile on at least one of the five scores.

Insert Table 2

Examples will be provided to facilitate the interpretation of Table 2. For example, in Chile, 61.9% of the sample have at least 1 of the 5 scores below the 25th percentile, 42.1% below the 16th percentile, 30.6% below the 10th percentile, 20.9% below the 5th percentile, and 12.1% below the 2nd percentile. These results are represented in Figure 1.

Insert Figure 1

In another example, in Spain, 54.7% of the sample have at least 1 of the 5 scores below the 25th percentile, 41.4% below the 16th percentile, 32.0% below the 10th percentile, 20.5% below the 5th percentile, and 11.3% below the 2nd percentile. The same 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, except for Cuba at the 25th percentile. The unusual numbers of low scores are four or more for the 25th percentile (three or more for Cuba), three or more for the 16th and 10th percentiles, and two or more for the 5th and 2nd percentiles. See Table 2.

Discussion

The availability of normative data for interpretation of performance on neuropsychological tests has been identified by practitioners as one of the main problems in neuropsychological assessment for Spanish-speaking countries (Arango-Lasprilla, Stevens, Morlett Paredes, Ardila & Rivera, 2017c; Rivera & Arango-Lasprilla, 2017). A

recent study addressed this issue and provided normative data for 10 neuropsychological tests to be used with children and adolescents from Chile, Colombia, Cuba, Ecuador, Guatemala, Honduras, Mexico, Paraguay, Peru, Puerto Rico, and Spain (Arango-Lasprilla, Rivera, & Olabarrieta-Landa, 2017b; Rivera & Arango-Lasprilla, 2017). Since the probability of obtaining one or more low scores is related to the number of measures in the battery, the aim of the present study was to analyze the base rate of low scores when several memory tests are administered, as memory impairments are frequent in pediatric populations with neurological diseases (Menlove & Reilly, 2015; Rayner et al., 2016; Serra-Grabulosa, 2005). Knowing the base rate of low scores in batteries with several measures will avoid false positive classifications of memory impairments in children. To our knowledge, this is the first study assessing the base rate of low scores on memory tests in normative samples from Latin America and Spain.

The results of this study show that obtaining one or more low scores is common in healthy children and adolescents when several memory tests are administered, in line with previous works on normative samples of healthy children (Brooks, Iverson, Sherman & Holdnack, 2009), healthy adults (Binder et all., 2009; Brooks & Iverson, 2010), healthy older adults (Palmer, 1998), and older adults with mild cognitive impairment (Oltra-Cucarella et al., 2018). Relatedly, the number of low scores decreased as the cut-off point to define unusual decrease, with fewer numbers of low scores for the 16th percentile compared to the 2nd percentile. When stratified by country, more than 37.4% (Peru) of children and adolescents obtained one or more scores below the 16th percentile, and between 10.7% (Honduras) and 16.1% (Cuba) obtained one or more scores below the 2th percentile. These results clearly show that the interpretation of performance on neuropsychological tests must take account of the number of tests administered and the

number of measures analyzed, as the distribution of frequencies of low scores differ for one single measure and for a battery with several measures.

Although less than 16% are expected to obtain a score one standard deviation below the mean for any single test, the number of measures below the 16th percentile obtained by less than 16% of the sample is three or more when 5 measures are analyzed. These results are in line with the works by Palmer (1998) in a sample of healthy older adults and by Brooks et al. (2009) in a sample of children and adolescents. Palmer (1998) analyzed the base rate of low scores among 5 memory measures in 152 healthy adults and found that the number of measures at or below the 10th percentile obtained by fewer than 10% of the sample was three or more. Brooks et al. (2009) analyzed the base rate of low scores among 6 measures in 1,000 healthy children and adolescents that form the normative sample for the Children's Memory Scale and found that approximately 10% of the sample obtained three or more scores at or below the 16th percentile. Our results showed that three or more scores at or below the 16th percentile were obtained by fewer than 10% of the healthy sample when five measures of verbal or visual memory were administered, which suggests that the use of one single low score to define abnormality should be abandoned.

In line with previous work, some researchers have found differences in low scores between different age groups (Brooks et al., 2009). However, in this study we adjusted for age, MLEP and sex and we found that even after controlling for these variables the distribution of the number of low scores change according to the percentile. These results support the notion that performance on neuropsychological tests must be compared to appropriate normative data and highlights the value of the multi-country normative data study for Spanish speaking individuals (Arango-Lasprilla et al., 2017b; Rivera & Arango-Lasprilla, 2017).

One of the previous findings that we could not replicate is the influence of intelligence estimates on the base rate of low scores. Brooks et al. (2010) and Binder et al. (2009) reported that the number of low scores was inversely related to intelligence estimates in healthy older adults, with the number of low tests obtained by fewer than 10% of the sample being 8, 5, and 3 for individuals with low average, average, and high average intellectual abilities respectively. As per the characteristics of our sample, we approached this relation by analyzing the association between the base rate of low scores and the children' parents educational level, due to the established correlation between parental education and cognitive functioning in children in past research (Schoenberg, Lange, & Saklofske, 2007; Thomas, Sukumaran, Lukose, George, & Sarma, 2007; van der Sluis, Willemsen, de Geus, Boomsma, & Posthuma, 2008). Brooks and colleagues (2010) showed a lower base rate of low scores on the NEPSY-II on children whose parents had more years of education, although they analyzed a different number of measures in each age range, from 7 measures in children 3-4 years-old to 17 measures in children 7-16 years-old. These findings were then replicated by Brooks (2011) in the Wechsler Intelligence Scale for Children – 4th Edition. One possible reason for this discrepancy is the difference in the range of years of education. While Brooks (2011) and Brooks et al. (2010) used ≤ 11, 12, 13-15 and 16+ categories, we dichotomized years of education as ≤ 12 or > 12, which restricts the range of possible numbers of low scores. For these reasons, as recommended by Brooks (2011) and Brooks et al. (2010), tables (i.e., stratified by parental education) should be used when children's cognitive functioning is to be compared to that of a sample of healthy individuals.

This study is not without limitations. First, these data apply only to children and adolescents aged between 6 and 17 years-old, enrolled in a regular private or public school, and with average or better intellectual functioning. Thus, these data should not be

applied to interpret performance on neuropsychological tests in individuals younger than 6 or older than 17, nor in illiterate children. Future studies will uncover whether the number of low tests differs between distinct levels of intellectual functioning in these Spanish-speaking samples.

Second, the results reported here apply only to batteries including five measures from the ROCF and the TAMV-I. Since the number of low scores is expected to increase as the number of measures in the battery increases (Binder et al., 2009; Brooks & Iverson, 2010; Brooks, 2011; Brooks et al., 2009), the base rate of low scores reported here should not be applied to batteries with a different number and type of measures.

Third, in this study MLPE was categorized as 12 or >12 years of education, because there were not enough subjects at each education year to allow the use of this variable as continuous variable in the analyses. As such, future studies should consider to using larger samples that allow the inclusion of MLPE as a continuous variable in the analyses.

In conclusion, these results show that, as it occurs in samples of healthy adults, cognitive deficits identifying in individual cases of children with suspected psychiatric or neurological diseases (e.g. epilepsy, attention/hyperactivity disorder, TBI) should not rely exclusively on the finding of one single low score, irrespective of the cut-off point used to define low scores. Taking account of normal variability in cognitive performance by giving allowance of some low scores may identify with greater certainty those children with true cognitive impairment who should be included in clinical trials for cognitive, academic, or behavioral disturbances. Practitioners from several Latin American countries and Spain can use the data reported here to minimize the number of children and adolescents erroneously classified as showing memory impairment when several tests

providing several measures are administered as part of the neuropsychological assessment.

Disclosure Statement

Funding Acknowledgements

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., Ertl, M. M., Muñoz Mancilla, J. M., García-Guerrero, C. E., Rodriguez-Irizarry, W., ... & Yacelga Ponce, T. P. (2017a). Rey–Osterrieth Complex Figure–copy and immediate recall (3 minutes): Normative data for Spanish-speaking pediatric populations. *NeuroRehabilitation, 41*(3), 593-603. doi:10.3233/nre-172241

Arango-Lasprilla, J. C., Rivera, D., & Olabarrieta-Landa, L. (Eds.). (2017b). *Neuropsicología infantil*. Manual Moderno.

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

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

Brooks, B. L. (2011). A study of low scores in Canadian children and adolescents on the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV). *Child Neuropsychology*, *17*(3), 281-289. doi:10.1080/09297049.2010.537255

Brooks, B., & Iverson, G. (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. doi:10.1093/arclin/acp100

Brooks, B. L., Iverson, G. L., Sherman, E. M. S., & Holdnack, J. A. (2009). Healthy children and adolescents obtain some low scores across a battery of memory tests. *Journal of the International Neuropsychological Society*, *15*(4), 613. doi:10.1017/S1355617709090651

Brooks, B. L., Sherman, E. M. S., & 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. doi:10.1093/arclin/acq005

Brown, L., Sherbenou, R. J. & Johnsen, S. K. (2009). *Test de Inteligencia No Verbal TONI-2*. Madrid, Spain: TEA Ediciones, S.A.

Bush, K., Kivlahan, D. R., McDonell, M. B., Fihn, S. D., & Bradley, K. A. (1998). The AUDIT Alcohol Consumption questions (AUDIT-C): An effective brief screening test for problem drinking. *Archives of Internal Medicine, 158*(16), 1789-1795. doi:10.1001/archinte.158.16.1789

Fastenau, P. S. (1996). Development and preliminary standardization of the "Extended Complex Figure Test" (ECFT). *Journal of Cinical and Experimental Neuropsychology, 18*(1), 63-76. doi:10.1080/01688639608408263

Folleco, J. A. E., Rivera, D., Acosta Barreto, M. R., Restrepo Botero, J. C., Díaz Camargo, E. A., Ramos Usuga, D., … & Arango-Lasprilla, J. C. (2017). *Datos normativos del Test de Copia y de Reproducción de Memoria de la Figura Geométrica Compleja de Rey en población Colombiana de 6 a 17 años de edad*. In J. C. Arango-Lasprilla, D. Rivera, & L. Olabarrieta-Landa (Eds). *Neuropsicología infantil* (pp. 197-206). Bogotá, Colombia: Manual Moderno.

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

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

Kovacs, M. (1992). *Children Depression Inventory CDI (Manual)*. Toronto, Canada: Multihealth.

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

Menlove, L., & Reilly, C. (2015). Memory in children with epilepsy: A systematic review. *Seizure*, *25*, 126-135. doi:10.1016/j.seizure.2014.10.002

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

Osterrieth, P. A. (1944). Le test de copie d'une figure complexe. *Archives de Psychologie, 30*, 206–356.

Palmer, B. (1998). Base rates of "impaired" neuropsychological test performance among healthy older adults. *Archives of Clinical Neuropsychology*, *13*(6), 503–511. doi:10.1016/S0887-6177(97)00037-1

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.

Rayner, G., Jackson, G. D., & Wilson, S. J. (2016). Mechanisms of memory impairment in epilepsy depend on age at disease onset. *Neurology*, *87*(16), 1642-1649. doi:10.1212/WNL.0000000000003231

Rey, A. (1941). L'examen psychologique dans les cas d'encéphalopathie traumatique (les problems). *Archives de Psychologie, 28*, 286–340.

Rey, A. (2009). *REY: Test de copia y de reproducción de memoria de figuras geométricas complejas.* Madrid, Spain: TEA Ediciones, S.A.

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

Rivera, D., Olabarrieta-Landa, L., & Arango-Lasprilla, J. C. (2017d). *Metodología para generar datos normativos en población colombiana de 6 a 17 años de edad*. In J.C. Arango-Lasprilla, D. Rivera, & L. Olabarrieta-Landa (Eds). *Neuropsicología infantil* (pp. 163-195). Bogotá, Colombia: Manual Moderno.

Rivera, D., Olabarrieta-Landa, L., & Arango-Lasprilla, J. C. (2017a). *Diseño y creación del Test de Aprendizaje y Memoria Verbal Infantil (TAMV-I) en población hispano hablante de 6 a 17 años de edad*. In J. C. Arango-Lasprilla, D. Rivera, & L. Olabarrieta-Landa (Eds.), *Neuropsicología infantil* (pp. 316-338). Bogotá, Colombia: Manual Moderno.

Rivera, D., Posada Abril, J. L., Benito-Sánchez, I., Acosta Barreto, M. R., Riaño Garzón, M. E., Herrera Bravo, M. A., … & Arango-Lasprilla, J. C. (2017b). *Datos normativos para el test de aprendizaje Verbal (TAMV-I) en población de 6 a 17 años de edad en Colombia*. In J. C. Arango-Lasprilla, D. Rivera, & L. Olabarrieta-

Landa (Eds.), *Neuropsicología infantil* (pp. 339-349). Bogotá, Colombia: Manual Moderno.

Rivera, D., Olabarrieta-Landa, L., Rabago Barajas, B. V., Irías Escher, M. J., Saracostti Schwartzman, M., Ferrer-Cascales, R., ... & Arango-Lasprilla, J. C. (2017c). Newly developed Learning and Verbal Memory Test (TAMV-I): Normative data for Spanish-speaking pediatric population. *NeuroRehabilitation, 41*(3), 695-706.

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. J., Testa, 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. doi:10.1017/s1355617708080387.

Serra-Grabulosa, J. M. (2005). Cerebral correlates of declarative memory dysfunctions in early traumatic brain injury. *Journal of Neurology, Neurosurgery & Psychiatry*, *76*(1), 129–131. doi:10.1136/jnnp.2004.027631

Shin, M. S., Park, S. Y., Park, S. R., Seol, S. H., & Kwon, J. S. (2006). Clinical and empirical applications of the Rey-Osterrieth Complex Figure Test. *Nature Protocols, 1*(2), 892-899. doi:10.1038/nprot.2006.115

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. doi:10.1111/j.1528-1167.2007.01376.x

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. *Behavior Genetics, 38*(4), 348-360. doi:10.1007/s10519-008-9212-5

Waber, D. P., & Holmes, J. M. (1985). Assessing children's copy productions of the Rey-Osterrieth Complex Figure. *Journal of Clinical and Experimental Neuropsychology, 7*(3), 264-280. doi:/10.1080/01688638508401259

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

	Sample	Age	Sex		Type of school		MLPE
	N	Mean (SD)	Girls	Boys	Public	Private	Mean (SD)
			n (%)	*n* (%)	*n* (%)	*n* (%)	
Chile	387	11.5 (3.5)	194 (50.1%)	193 (49.9%)	195 (50.4%)	192 (49.6%)	12.3 (3.0)
Colombia	1657	11.4 (3.3)	863 (52.1%)	794 (47.9%)	836 (50.5%)	821(49.5%)	12.3 (3.8)
Cuba	381	11.5 (3.5)	190 (49.9%)	191 (50.1%)	381 (100%)	0 (0.0%)	16.2 (1.8)
Ecuador	302	11.4 (3.5)	175 (57.9%)	127 (42.1%)	159 (52.6%)	143 (47.4%)	14.4 (3.6)
Guatemala	203	10.7 (2.5)	94 (46.5%)	108 (53.5%)	112 (55.2%)	91 (44.8%)	10.5 (4.1)
Honduras	300	11.2 (3.2)	161 (53.7%)	139 (46.3%)	155 (51.7%)	145 (48.3%)	12.8 (3.7)
Mexico	934	11.4 (3.5)	481 (51.5%)	453 (48.5%)	574 (61.5%)	360 (38.5%)	13.1 (3.9)
Paraguay	300	11.6 (3.5)	161 (53.7%)	139 (46.3%)	141 (47.0%)	159 (53.0%)	14.1 (2.9)
Peru	348	12.0 (3.3)	171 (49.1%)	177 (50.9%)	187 (53.7%)	161 (46.3%)	12.5 (2.4)
Puerto Rico	215	12.2 (3.6)	120 (55.8%)	95 (44.2%)	133 (61.9%)	82 (38.1%)	14.5 (2.6)
Spain	1003	11.3 (3.4)	518 (51.6%)	485 (48.4%)	546 (54.4%)	457 (45.6%)*	14.0 (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 learning and memory low scores below the specified percentile cutoff by country.

	Number of low scores	All countries	Chile	Colombia	Cuba	Ecuador	Guatemala	Honduras	Mexico	Paraguay	Peru	Puerto Rico	Spain
<25th percentile	None	42.9%	38.1%	41.9%	49.7%	43.4%	46.0%	39.9%	40.0%	40.4%	49.9%	40.9%	45.3%
	1+	57.1%	61.9%	58.1%	50.3%	56.6%	54.0%	60.1%	60.0%	59.6%	50.1%	59.1%	54.7%
	2+	31.8%	32.7%	32.7%	27.7%	32.8%	29.0%	31.6%	32.9%	34.7%	31.8%	38.1%	28.9%
	3+	12.4%	12.1%	12.9%	8.1%	13.6%	14.0%	12.4%	12.1%	17.5%	10.4%	14.2%	11.9%
	4+	3.7%	4.0%	4.3%	2.4%	5.0%	4.5%	3.4%	3.0%	5.7%	4.2%	2.3%	3.1%
	5	.9%	.8%	.9%	.3%	2.3%	.5%	.7%	.3%	2.0%	1.8%	--	.9%
<16th percentile	None	55.8%	57.9%	53.4%	60.5%	52.6%	57.0%	57.7%	54.2%	51.9%	62.6%	51.1%	58.6%
	1+	44.2%	42.1%	46.6%	39.5%	47.4%	43.0%	42.3%	45.8%	48.1%	37.4%	48.9%	41.4%
	2+	19.1%	18.5%	18.9%	16.1%	20.9%	18.0%	19.2%	20.3%	22.2%	17.2%	21.0%	18.6%
	3+	5.9%	7.0%	6.6%	4.3%	6.6%	7.5%	6.9%	5.2%	6.4%	5.6%	5.1%	4.6%
	4+	1.6%	1.9%	1.7%	1.9%	3.0%	2.0%	2.1%	.9%	1.0%	1.5%	.6%	1.4%
	5	.2%	.5%	.1%	--	1.0%	--	--	--	.3%	.9%	--	--
<10th percentile	None	67.0%	69.4%	67.1%	69.4%	64.9%	66.5%	67.4%	65.6%	63.0%	72.7%	60.8%	68.0%
	1+	33.0%	30.6%	32.9%	30.6%	35.1%	33.5%	32.6%	34.4%	37.0%	27.3%	39.2%	32.0%
	2+	12.0%	12.1%	11.5%	12.9%	13.9%	10.5%	11.3%	12.2%	14.1%	11.6%	14.2%	10.9%
	3+	3.3%	3.2%	3.2%	3.2%	4.0%	4.5%	4.1%	2.8%	3.0%	2.7%	2.8%	2.9%
	4+	.8%	1.1%	.8%	1.1%	1.3%	.5%	1.0%	.2%	.3%	.9%	--	.4%
	5	.1%	--	.1%	--	.3%	--	--	--	.3%	--	--	--
<5th percentile	None	78.3%	79.1%	77.6%	78.8%	79.1%	76.0%	78.7%	78.5%	77.1%	80.1%	72.7%	79.5%
	1+	21.7%	20.9%	22.4%	21.2%	20.9%	24.0%	21.3%	21.5%	22.9%	19.9%	27.3%	20.5%
	2+	6.3%	5.9%	6.3%	7.5%	6.3%	4.5%	5.2%	5.9%	7.1%	7.7%	4.5%	6.5%
	3+	1.3%	1.6%	1.4%	1.3%	1.7%	2.0%	1.0%	.9%	.7%	1.5%	--	1.4%
	4+	.2%	.3%	.3%	--	.3%	--	--	--	.3%	--	--	.1%

	5	--	--	--	--	--	--	--	--	.3%	--	--	--
<2th percentile	None	87.4%	87.9%	87.2%	83.9%	88.1%	86.5%	89.3%	87.9%	88.2%	85.5%	84.1%	88.7%
	1+	12.6%	12.1%	12.8%	16.1%	11.9%	13.5%	10.7%	12.1%	11.8%	14.5%	15.9%	11.3%
	2+	2.7%	2.4%	2.2%	4.6%	3.0%	3.5%	2.4%	1.9%	1.3%	5.9%	1.1%	3.0%
	3+	.4%	.5%	.4%	--	.3%	.5%	.7%	--	--	.9%	--	.5%
	4+	.1%	--	.2%	--	--	--	--	--	--	--	--	--
	5	--	--	--	--	--	--	--	--	--	--	--	--

Figure 1. Cumulative proportion of Chilean children with the specified number of adjusted learning and memory low scores below the specified percentile cutoff.
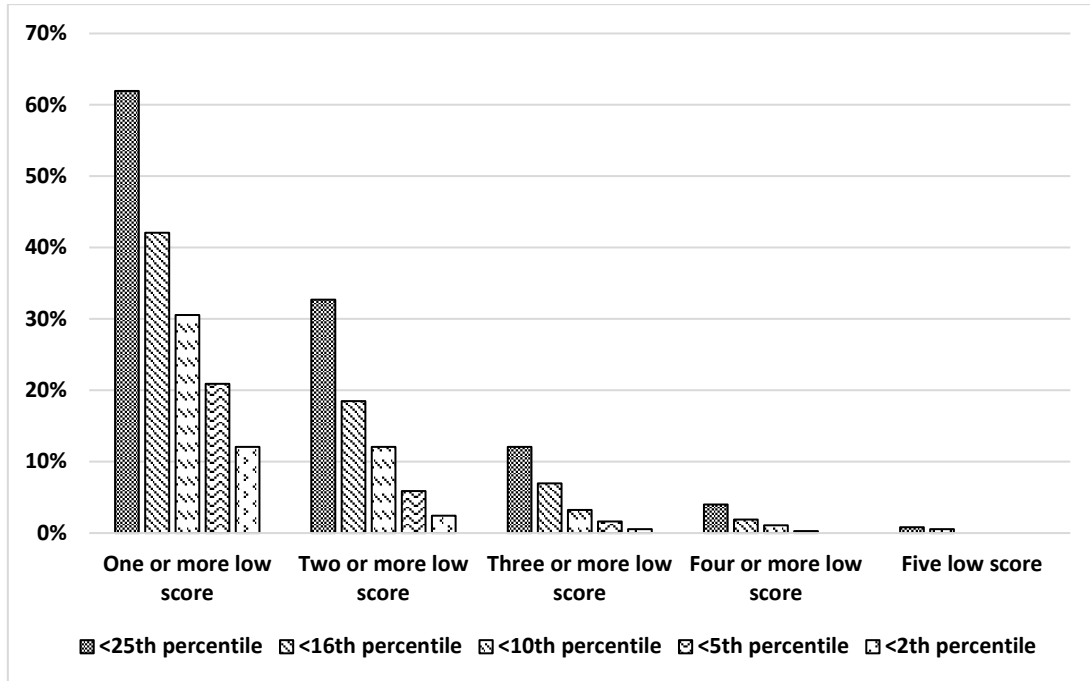
Figure 2. Cumulative proportion of Spanish children with the specified number of adjusted learning and memory low scores below the specified percentile cutoff.