15-item version of the Boston Naming Test: Normative data for the Latin-American Spanish-speaking adult population

Ana delCacho-Tena, MSc.
Department of Health Science, Public University of Navarre, Pamplona, Spain
delcacho.151053@e.unavarra.es
ORCID: 0000-0003-0649-0226

Kritzia Merced, Ph.D.
Virginia Commonwealth University, Richmond, VA, USA
Central Virginia Veterans Affairs Health Care System, Richmond, VA, USA
kmerced.morales@gmail.com

Paul B. Perrin, Ph.D.
School of Data Science and Department of Psychology, University of Virginia, Charlottesville, VA, US.
perrin@virginia.edu

Juan Carlos Arango-Lasprilla, Ph.D.
Giunti Psychometrics, Italy
Department of Psychology. Virginia Commonwealth University. Richmond. USA.
jcalasprilla@gmail.com
ORCID: **0000-0002-7184-8311**

Laiene Olabarrieta-Landa, Ph.D.*
Department of Health Science, Public University of Navarre, Pamplona, Spain
Instituto de Investigación Sanitaria de Navarra (IdiSNA), Pamplona, Spain
laiene.olabarrieta@unavarra.es
ORCID: 0000-0002-8305-8720

Diego Rivera, M. P. H., Ph.D.*
Department of Health Science, Public University of Navarre, Pamplona, Spain
Instituto de Investigación Sanitaria de Navarra (IdiSNA), Pamplona, Spain
diego.rivera@unavarra.es
ORCID: 0000-0001-7477-1893

*Equal last-author contribution

Corresponding Author:
Diego Rivera, M. P. H., Ph.D.
Department of Health Sciences
Public University of Navarre
Arrosadia Campus, s/n. 31006. Pamplona. Spain
diego.rivera@unavarra.es

Abstract

Objective: Naming is commonly impaired in people with neurodegenerative diseases and brain injury, and as a result, its accurate assessment is essential. The aim of this study was to provide normative data for the 15-item Spanish version of the Boston Naming Test (BNT) for an adult population from eight Latin American countries.

Method: The total sample consisted of 2,828 participants from Argentina, Chile, Cuba, El Salvador, Mexico, Honduras, Paraguay, and Puerto Rico. Multiple regressions were used to provide normative data following a four-step process.

Results: Regression models showed a linear effect of age on the 15-item BNT total score for Argentina, Chile, Mexico, and Puerto Rico. In contrast, Paraguay showed a quadratic age effect. Almost all countries showed a positive linear effect of education, except Cuba which presented a quadratic effect. Sex was a significant predictor in Argentina, Chile, Cuba and Mexico, where women scored lower than men.

Conclusions: This study provides normative data for the 15-item Spanish version of the BNT and offers a free open-access calculator that will assist in the efficacious use of this neuropsychological test in clinical practice and research in the future.

Keywords: 15-item Boston Naming Test, Neuropsychological assessment, Latin American adult population, Healthy people, Normative Data.

Key Points

Question: To provide normative data for the 15-item Boston Naming Test (BNT) for adults from eight Latin American countries.

Findings: Regression models showed a linear or quadratic effect of age, education and sex around the countries on the 15-item BNT.

Importance: This study provides normative data for the 15-item Spanish version of the BNT and offers a free open-access calculator.

Next steps: This study will assist in the efficacious use of this neuropsychological test in clinical practice and research in the future.

Introduction

In recent decades, the prevalence of people with neurodegenerative disorders such as dementia, multiple sclerosis (MS), or brain injury (e.g., traumatic brain injury [TBI]) has increased (e.g. Alegret et al., 2012; King et al., 2020; Lo et al., 2021). Therefore, accurate assessment and treatment of these diseases and disorders have become a priority for the vast majority of governments and health systems worldwide (e.g. Kristinsson et al., 2020; Mueller et al., 2018). Among the cognitive impairments that people with different neurological disorders usually present, the most frequent are attention impairments (Boerma et al., 2017), memory deficits (Blom et al., 2021), difficulties with executive functions (Fisher et al., 2018), and language impairment (Blom et al., 2021).

Language encompasses a series of processes such as repetition (Kristinsson et al., 2020), fluency (Botha & Josephs, 2019; Kristinsson et al., 2020; Mueller et al., 2018), comprehension (Kristinsson et al., 2020; Mueller et al., 2018), naming (Kristinsson et al., 2020), reading (Tabet et al., 2020), and writing (Berger et al., 2005). It allows humans to express ideas and emotions, as well as to understand what others think and feel. Language is commonly impaired in people with brain damage and neurodegenerative diseases, and its accurate assessment is essential to determine which of the underlying language processes are impaired and may need targeted intervention (Alegret et al., 2012; Botha et al., 2019; Kristinsson et al., 2020). In particular, naming is the most commonly impaired language process among these individuals, and its evaluation is important because it is sensitive to changes in brain function (Kristinsson et al., 2020; Mueller et al., 2018; Roth et al., 2018). Different instruments measure this language process, but one of the most widely used tests around the world is the standard-Boston Naming Test (BNT; Leite et al., 2017).

The BNT (Kaplan et al., 1983) consists of 60 drawings presented in order of difficulty. A total score is based on spontaneous correct answers and correct responses after providing a semantic cue. The BNT evaluates the ability to access the lexicon and naming, and it has been widely used in people with different types of diseases/disorders, such as dementia (Clark et al., 2009; Gardini et al., 2013;), TBI (Kavé et al., 2011; Zgaljardic et al., 2013), MS (Rivera et al., 2022; Roth et al., 2018), Parkinson Disease (e.g. Henry & Crawford, 2004), and stroke (Brady et al., 2001; Kim et al., 2011), among others. Normative data have been developed for the BNT in different countries including Turkey (Ekinci-Soylu & Cangöz., 2018), the Netherlands (Van der Elst et al., 2006), Spain (Peña-Casanova et al., 2009; Rami et al., 2008), Portugal (Vicente et al., 2021), and several Latin American countries (Kim et al., 2018; Olabarrieta-Landa et al., 2015; Rodríguez-Lorenzana et al., 2020).

Performance on the BNT has been associated with age (Gollan et al., 2007; Olabarrieta-Landa et al., 2015; Peña-Casanova et al., 2009; Rodríguez-Lorenzana et al., 2020; Soylu & Cangöz., 2018), gender (Olabarrieta-Landa et al., 2015; Peña-Casanova et al., 2009; Rodríguez-Lorenzana et al., 2020; Soylu & Cangöz., 2018), education (Gollan et al., 2007; Olabarrieta-Landa et al., 2015; Peña-Casanova et al., 2009; Rodríguez-Lorenzana et al., 2020; Soylu & Cangöz., 2018), and bilingualism (Erdodi et al., 2017; Gollan et al., 2007), so that the score on the standard-BNT decreases with age (Olabarrieta-Landa et al., 2015; Peña-Casanova et al., 2009) and increases respect more years of education (Olabarrieta-Landa et al., 2015; Peña-Casanova et al., 2009). On the other hand, bilingual individuals score better on the standard-BNT (e.g., Gollan et al., 2007), while no gender differences are reported among the literature (Olabarrieta-Landa et al., 2015; Peña-Casanova et al., 2009).

Different short versions of the BNT exist including the 30-item BNT version (Graves et al., 2004; Leite et al., 2017; Saxton et al., 2000) and the 15-item BNT version (Attridge et al., 2020; Graves et al., 2004; Leite et al., 2017), both showing appropriate representation of the standard-BNT (Erdodi et al., 2018; Hobson et al., 2011; Leite et al., 2017). In the English version, the correlation between the standard-BNT and the 30-item BNT is 0.98 and for the

15-item BNT is 0.93 (Graves et al., 2004), while in the Spanish version, the correlation between de standard BNT and the 15-item version ranges between 0.82 and 0.93 (Nebreda et al., 2011). No articles have been found on the correlation of 30-item BNT and standard-BNT in the Spanish version. The main advantage of these shorter versions is the ability to include them in different neuropsychological batteries as they require less administration and scoring time (Alegret et al., 2012; Zgaljardic et al., 2013). However, there is a lack of normative data for the shorter versions of the BNT. Although there are no studies that reports using BNT normative data from Anglo-Saxon populations shows impairment in Latino population, there are studies that state that between 28% and 57% of the clinicians use norms from other countries with their Latinos patients (Arango-Lasprilla et al., 2017). Additionally, Kim et al. (2018) showed that the standard-BNT administered to bilingual people both in English and Spanish had no equivalence on the execution.

In light of this, it is argued that having norms for a short version of the BNT can be useful for clinicians, as it may mitigate time and demand constraints, making it easier to include in testing protocols. Thus, the aim of this study was to provide normative data for the 15-item version of the BNT for an adult population in eight Latin-American countries.

Material and Methods

Participants

The sample consisted of 2,828 healthy participants from eight Latin American countries: Argentina (*n* = 316), Chile (*n* = 245), Cuba (*n* = 292), El Salvador (*n* = 244), Honduras (*n* = 181), Mexico (*n* = 1129), Paraguay (*n* = 258), and Puerto Rico (*n* = 163). The sampling strategy resulted from considering the age distribution, education level, and sex of participants by country. The demographic characteristics by country can be found in Table 1.

Insert Table 1

The inclusion criteria were being: (a) between 18 to 95 years of age, (b) born and currently living in the country where the protocol was conducted, (c) native Spanish-speaker, (d) having at least 1 year of formal education, (e) able to read and write at the time of evaluation, (f) having a score of 23 or more on the Mini-Mental State Examination (MMSE; Folstein et al., 1975; Villaseñor-Cabrera et al., 2010; Ostrosky-Solís et al., 2000), (g) having a score of 4 or less on the Patient Health Questionnaire–9 (PHQ-9), which is a 9-item psychological tool that assesses the frequency of depressive symptoms during the last two weeks on a 0 to 3 scale, where 0 means “not at all” and 3 means “almost every day” (Kroenke et al., 2001), and (h) having a score of 90 or more on the Barthel Index, a 10-item scale that evaluates basic activities of daily living (ADL), where a score of 0 means the participant needs help in all ADL and 100 means that the person is completely independent (Mahoney & Barthel, 1965; Zhang et al., 2022).

A self-report questionnaire was administered to collect data about medical history and health status. People who met the following criteria were excluded from the study: (a) having a neurological diagnosis or psychiatric condition, (b) having consumed and/or daily use of an illicit substance, (c) having a history of non-controlled systemic disease (e.g., diabetes mellitus), (d) having a regular use of pain drugs or other medications that may impact cognitive functioning, and/or (e) having a severe visual and/or hearing deficit. Participants were volunteers and did not receive financial compensation for their involvement in the study. For further information about data collection see Guàrdia-Olmos et al. (2015), and Olabarrieta-Landa et al. (2015).

Procedure

This study was conducted as part of a larger study to generate normative data in healthy adults from several Latin American countries. As the coordinating institution, The

'Masked' Ethics Committee approved this study. Data collection was conducted between March 2013 and August 2014. Participants were recruited through advertisements in different institutions (local business, community facilities and universities). All participants were volunteers and contacted by a member of the team to explain the characteristics of the study, sign an informed consent form and collect sociodemographic information of interest. The 15-item Spanish version of the BNT was applied in conjunction with other tests in a single day, and the battery had a total duration of 70 minutes.

Measure

The 15-item Spanish version of the BNT used in this study was created by Kaplan et al. (2005). Participants had to name a set of black and white illustrations, which were presented from least to most difficult. The administration of the short-BNT was as follows: (a) the slides were presented in order one at a time, (b) the response had to be given within 20 seconds, (c) if participants said a word different from the expected one, it was transcribed and a semantic cue was given only if the incorrect response was not within the correct semantic category, (d) if the semantic cue was not sufficient, a phonological cue was provided. The total score for each subject was calculated considering the number of correct spontaneous answers and the number of correct answers after a semantic clue (Chen et al., 2014; Olabarrieta-Landa, et al., 2015; Tamaš et al., 2021). Thus, synonyms were not considered correct responses. The names of the items used in the Spanish version of the BNT were listed in supplemental material A1.

Statistical Analyses

Exploratory Data Analysis

Spearman correlations (r_S) were calculated between the 15-item BNT total score and age, education, and sex (see Table 2). To determine whether significant differences between the countries existed on the 15-item BNT, an analysis of covariance (ANCOVA) was performed, where age and education were covariates and country as fixed-effect factor.

The Effects of Sociodemographic Variables and Normative Data

The effects of demographic variables on the 15-item-BNT total score were analyzed by multiple regression models in each country. The full-regression model included the following predictors: age, age^2, education, education^2, sex, and the two-level interactions between these variables ($Interactions_i$). The predictors age^2 and education^2 were the result of the age and education centered, respectively, using the expression: (age – mean age; education – mean education)^2. This form of centering predictors has been commonly used to avoid multicollinearity (Kutner et al., 2005). Squared predictors were used in the full regression models to allow for quadratic effects of these variables on the 15-item BNT total score. Additionally, sex was dummy coded as woman = 0 and man = 1. The multiple regression model computed for each country was:

$$\hat{Y}_i = \mathrm{B}_0 + \mathrm{B}_i\left(age - \bar{x}_{age\ by\ country}\right)_i + \mathrm{B}_2\left(age - \bar{x}_{age\ by\ country}\right)_i^2 + \mathrm{B}_3\left(education - \bar{x}_{education\ by\ country}\right)_i + \mathrm{B}_4\left(education - \bar{x}_{education\ by\ country}\right)_i^2 + \mathrm{B}_5 sex_i + \mathrm{B}_k Interactions_i + \varepsilon_i.$$

The model assumes that the residuals ε_i are normally distributed, with means 0 and standard deviation σ_ε^2 ($\varepsilon_i \sim N[0, \sigma_\varepsilon^2]$). When creating the regression models for each country, all predictors and their respective two-way interactions were used following the methodology of Rivera et al. (2019). The independent variables that were not significant for each country were removed, and the regression model was adjusted again with the remaining parameters. A Bonferroni alpha-level of .005 (.05/10 $\approx$ $number\ of\ predictor$s) was used to avoid Type I errors due to multiple testing. Four assumptions were analyzed in this study for each regression model: (a) multicollinearity, which was evaluated with the Variance Inflation Factor (VIF) that be no greater than 10 (Kutner et al., 2005), (b) homoscedasticity, by applying Levene's test on the residuals of the predicted scores for each country (Van der Elst et al., 2006), (c) standardized residual normality applying the Kolmogorov-Smirnov test, and (d) the existence of values influencing the predictors or the total score, calculated by Cook's distance, whose distribution is characterized by $F(p, n - p)$. In this mathematical expression, p is the number of the total predictors in each final regression model (including the constant) and n is the sample size (Kutner et al., 2005).

Using the final regression models, normative data adjusted for the appropriate sociodemographic variables were generated using a four-step methodology (Rivera et al., 2019, 2020; Van der Elst et al., 2006): (a) calculation of the predicted test score ($\hat{Y}_i$), which is based on the estimate of the fixed effects independent variables (parameters) of the final regression model obtained; (b) computation of the residual value (e_i), using the difference between the raw score of the test and the $\hat{Y}_i$, expressed as $e_i = Y_i - \hat{Y}_i$; (c) making use of the residual standard deviation (SD_e) values, calculated from the final regression model as follows: $z_i = e_i/SD_e$; and finally (d) the percentile of each z_i calculated in the previous step was obtained using the cumulative distribution function, (only if the assumption of normality of the residuals in the sample was met) or by means of the empirical-type distributive function, in case this assumption was not met. This four-step process was applied to the total score by each country. All analyses were performed using R version 4.0.3 for macOS (R-Development Core Team, 2020).

Results

Exploratory Data Analysis

Spearman's correlation coefficients (r_s; Table 2, Supplemental material A2, and Supplemental material A3) showed that age was significantly associated with the 15-item BNT total score for Chile (r_s = -.32; p<.0001; r^2= .10), Cuba (r_s = -.23; p<.0001; r^2=.06), El Salvador (r_s = -.19; p<.01; r^2=.04), Mexico (r_s = -.35; p<.0001; r^2=.12), Paraguay (r_s = -.40;

p<.0001; r^2=.16), and Puerto Rico (r_s = -.24; *p*<.01; r^2=.06). In addition, education showed a significant relationship for all countries in all Spearman's correlations (ρ's >26; *p*<.0001). Sex was significant for Argentina (r_s = .16; *p*<.01; r^2=.02), Chile (r_s = .19; *p*<.01; r^2=.03), Cuba (r_s = .27; *p*<.0001; r^2=.07), and Mexico (r_s = .16; *p*<.0001; r^2=.03; Table 2). An ANCOVA showed a significant difference on the 15-item BNT total score across countries (*F*= 13.48; *p*<.0001; partial η^2= 0.13).

Insert Table 2

Model Assumptions and the Effects of Sociodemographic Variables

The four assumptions of the multiple linear regression analysis were met in most of the countries. Regarding multicollinearity, all the countries presented values of VIF <2, so that multicollinearity was not observed in any country. The maximum Cook's distance value was .097 for Chile, indicative of the existence of influential cases (relating this value to an *F*(4, 241) distribution yields a percentile value of 1.7, which is well below the threshold percentile value = 50).With respect to homoscedasticity, Levene's test indicated that the final models in some countries had heteroscedasticity, except Argentina, Chile, Honduras, and Puerto Rico. In case of heteroscedasticity, SD_e were calculated for each quartile predicted-score group (see Table 3, last column). The analysis of the standardized residual values of the regression models showed that they were normally distributed (evaluated with the Kolmogorov-Smirnov test), except for Paraguay, where the empirical distribution was used to convert z_i.to percentile.

Insert Table 3

The final regression models showed a linear effect of age on the 15-item BNT total scores for Argentina, Chile, Mexico, and Puerto Rico, where BNT total scores decreased with age. However, a quadratic age effect on the 15-item BNT total score was observed in Paraguay, where the 15-item BNT showed the higher test-score between 18 and 20 years of age, and decreased onwards in a curvilinear pattern. In relation to education, almost all countries showed a positive linear effect, such that greater years of education led to a higher 15-item BNT total score. Nonetheless, Cuba, had a quadratic effect of education. As a result, the mean predicted score showed an increment of the test-scores until 13 and 14 years of education, then the pattern tended to decrease slowly. Sex was significant in Argentina, Chile, Cuba, and Mexico, such that women scored lower than men (see Table 3).

Calculation of Normative Data

The procedure for transforming raw scores into percentiles for the 15-item Spanish version of the BNT total score in each country was completed using the four-step procedure described in the statistical analyses section. A brief example of the calculation of normative data follows. To calculate the percentile for a 30-year-old Chilean male, with 8 years of education and who scored a 13 on the 15-item Spanish version of the BNT total score, the following steps should be taken: (a) to calculate the expected 15-item BNT test score for this participant: which provides the final regression models in Chile use the "b" regression weights of each predictor to estimate the expected total score which equals: $\hat{Y}_i = 10.76 + (-0.03 * (30 - 57.4)_i) + 0.21 * (8 - 9.55)_i + 0.86 * 1$= 12.117. (b) Calculate residual value (e_i), that is, $e_i = 13 - 12.117 = 0.884$. (c) The residual is standardized by means of the SD_e (residual) value of the model (see column in Table 3 to obtain the country-specific SD_e value according to the participant's $\hat{Y}$). In this case $z_i = \frac{e_i}{SD_e} = \frac{0.884}{0.863} = 1.024$.

To avoid human errors when calculating the normative data, and to facilitate the process to clinicians, a calculator was created in Microsoft Excel, and a clinician must enter the following information: country, age, education, sex of participant, and the total score in the 15-item BNT. This calculator is freely available for all users and may be downloaded at https://neuropsychologylearning.com/datos-normativos-archivos-descargables/

Discussion

The goal of the current study was to validate the 15-item version of the BNT and to provide normative data for the adult population in eight Latin American countries. The data were explored through country-specific adjustments for sex, education, and age in each model. Out of the eight countries, a weak but significant positive correlation was found between sex and total scores in Argentina, Chile, and Cuba - with Mexico showing a slightly stronger correlation. When included in the regression models, sex remained significant in those four countries, showing slightly higher scores in men compared to women. The minimal correlation between sex and 15 item-BNT total scores is broadly consistent with other studies examining the BNT standard or short forms that have found no or minimal sex differences (Abeare et al., 2022; Fastenau et al., 1998; Fillenbaum et al., 1997; Mitrushina et al., 2005; Olabarrieta-Landa et al., 2015). Moreover, in the four other countries (i.e., Puerto Rico, Paraguay, Honduras, and El Salvador) showing no differences by sex, there may be specific cultural or gender norms impacting the scores. Plausible explanations for these gender differences, or lack thereof, may be the influence of the educational system structure and sex enrollment by country, as well as a reflection of the items being more aligned with traditional male roles (Aranciva et al., 2012).

All countries in this study showed a positive significant correlation (low to moderate in strength) between education and the 15-item BNT total scores. The models showed that 15-item BNT total scores increased with greater years of education across all countries, thus, adding to previous research obtaining similar effects (Abeare et al., 2022; Fernández-Blázquez & Ruiz-Sánchez León, n.d.; Fillenbaum et al., 1997; Olabarrieta-Landa et al., 2015).

Interestingly, education was the only factors that was significant across all countries, further suggesting that having greater years of education consistently shows more language skills irrespective of educational structure or country of residence. Accordingly, these findings should be leveraged in the use of education-adjusted norms generated for each country when administering the 15-item BNT short form. With this, neuropsychologists in Latin American will not only be able to use a shorter version of the BNT, but they will also be also to obtain a score accounting for differences in each country's educational structures.

In relation to age, 15-item BNT total scores showed significant but small associations in four out of the eight countries, including Argentina, Cuba, Mexico, and Paraguay, such that total scores decreased with advancing age. The findings are similar to previous studies that have shown lower BNT total scores in older adults (Abeare et al., 2022; Fernández-Blázquez et al., 2012; Fillenbaum et al., 1997; Neils et al., 1995). Moreover, these findings are also congruent with the results in a study by Olabarrieta and colleagues (2015) examining the standard-BNT utilizing this same sample, but including two additional countries.

Constraints on generality

This study has several limitations that should be considered in the context of future clinical research. First, the study examined the 15-item Spanish version of the BNT in eight countries, and thus, generalizability to other countries is limited, and caution should be exerted when assessing a participant from a country where data were not collected. As such, a potential area for future research is examining generalizability in other Spanish-speaking Latin-American countries including Ecuador, Panama, and the Dominican Republic, among others. Second, participants in the current study were native Spanish-speakers, and data were not collected regarding whether participants spoke additional languages or local dialects. It is possible that speaking a second language or local dialect has effects on a participant's BNT performance. Accordingly, future research may focus on examining the influence of monolinguism versus bilingualism on BNT performance. Relatedly, the third point considers sample limitations. The sample was a convenience sample with specific inclusion criteria that was collected in specific regions of the countries in the study. This strategy may have excluded large segments of individuals that otherwise would be present in a national sample. For instance, participants in this study were required to have the ability to read and write and did not have a history of neurological conditions. Respectively, to expand on the current findings, future studies should examine the 15-item BNT's performance in individuals who are illiterate, have a history of neurological conditions, and children, as well as using a nationally representative sample in these countries.

Conclusion

Despite these limitations, the current study is the largest neuropsychological normative study in Latin America for norming the 15-item version of the BNT among Spanish-speakers and represents an initial step toward a larger and more representative normative study in the future. This study generated normative data from eight counties with a sizable sample, and it will assist in the efficacious use of this neuropsychological test in clinical and research practice in the future.

References

Abeare, K., Cutler, L., An, K. Y., Razvi, P., Holcomb, M., & Erdodi, L. A. (2022). BNT-15: Revised Performance Validity Cutoffs and Proposed Clinical Classification Ranges. *Cognitive and behavioral neurology: official journal of the Society for Behavioral and Cognitive Neurology*, *35*(3), 155–168. https://doi.org/10.1097/WNN.0000000000000304

Alegret, M., Espinosa, A., Vinyes-Junqué, G., Valero, S., Hernández, I., Tárraga, L., Becker, J. T., & Boada, M. (2012). Normative data of a brief neuropsychological battery for Spanish individuals older than 49. *Journal of clinical and experimental neuropsychology*, *34*(2), 209–219. https://doi.org/10.1080/13803395.2011.630652

Aranciva, F., Casals-Coll, M., Sánchez-Benavides, G., Quintana, M., Manero, R. M., Rognoni, T., Calvo, L., Palomo, R., Tamayo, F., & Peña-Casanova, J. (2012). Estudios normativos españoles en población adulta joven (Proyecto NEURONORMA jóvenes): normas para el Boston Naming Test y el Token Test. *Neurología*, *27*(7), 394–399. https://doi.org/10.1016/j.nrl.2011.12.016

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

Attridge, J., Zimmerman, D., Rolin, S., & Davis, J. J. (2020). Comparing Boston naming test short forms in a rehabilitation sample. *Applied Neuropsychology: Adult*. *25*, 1-6. https://doi.org/10.1080/23279095.2020.1811984

Berger, A.K., Fratiglioni, L., Winblad, B., Bäckman, L. (2005). Alzheimer's disease and depression: preclinical comorbidity effects on cognitive functioning. *Cortex*. *41*(4), 603-12. https://doi.org10.1016/s0010-9452(08)70200-4

Blom, E., Berke, R., Shaya, N., Adi-Japha, E. (2021). Cognitive flexibility in children with Developmental Language Disorder: Drawing of nonexistent objects. *J Commun Disord.;93*. https://doi.org/10.1016/j.jcomdis.2021.106137.

Boerma, T., Leseman, P., Wijnen, F., Blom, E. (2017) Language Proficiency and Sustained Attention in Monolingual and Bilingual Children with and without Language Impairment. *Front Psychol. 8*. https://doi.org/ 10.3389/fpsyg.2017.01241.

Botha, H., & Josephs, K. A. (2019). Primary Progressive Aphasias and Apraxia of Speech. *Continuum (Minneapolis, Minn.)*, *25*(1), 101–127. https://doi.org/10.1212/CON.0000000000000699

Brady, C. B., Spiro, A. III, McGlinchey-Berroth, R., Milberg, W., and Gaziano, J. M. (2001). Stroke risk predicts verbal fluency decline in healthy older men: evidence from the normative aging study. *J. Gerontol. Ser. B 56*, 340–346. https://doi.org/10.1093/geronb/56.6.P340

Chen, T. B., Lin, C. Y., Lin, K. N., Yeh, Y. C., Chen, W. T., Wang, K. S., & Wang, P. N. (2014). Culture qualitatively but not quantitatively influences performance in the Boston naming test in a chinese-speaking population. *Dementia and geriatric cognitive disorders extra*, *4*(1), 86–94. https://doi.org/10.1159/000360695

Clark, L. J., Gatz, M., Zheng, L., Chen, Y. L., McCleary, C., and Mack, W. J. (2009). Longitudinal verbal fluency in normal aging, preclinical, and prevalent Alzheimer's disease. *Am. J. Alzheimer's Dis. Other Dement. 24*, 461–468. https://doi.org/10.1177/1533317509345154

Erdodi, L. A., Jongsma, K. A., Issa, M. (2017). The 15-item version of the Boston Naming Test as an index of English proficiency. *Clin Neuropsychol.*, *31*(1), 168-178. https://doi.org/10.1080/13854046.2016.1224392

Erdodi, L. A., Dunn, A. G., Seke, K. R., Charron, C., McDermott, A., Enache, A., Maytham, C., Hurtubise, J. L. (2018) The Boston Naming Test as a Measure of Performance Validity. *Psychol. Inj. and Law 11*, 1–8. https://doi.org/10.1007/s12207-017-9309-3

Fastenau, P. S., Denburg, N. L., & Mauer, B. A. (1998). Parallel short forms for the Boston Naming Test: psychometric properties and norms for older adults. Journal of Clinical and Experimental *Neuropsychology, 20*(6), 828–834. https://doi.org/10.1076/jcen.20.6.828.1105

Fernández-Blázquez, & Ruiz-Sánchez León. (n.d.). Nueva versión reducida del test de denominación de Boston para mayores de 65 años: aproximación desde la teoría de respuesta al ítem. *Revista de Neurologia*. http://www.logicortex.com/wp-content/uploads/Art%C3%ADculo_31_Reducci%C3%B3n%C3%8DtemsBNT_RevNeurol_2012.pdf

Fillenbaum, G. G., Huber, M., & Taussig, I. M. (1997). Performance of elderly White and African American community residents on the abbreviated CERAD Boston Naming Test. *Journal of clinical and experimental neuropsychology, 19*(2), 204–210. https://doi.org/10.1080/01688639708403851

Fisher, E. L., Barton-Hulsey, A., Walters, C., Sevcik, R. A., Morris, R. (2019). Executive Functioning and Narrative Language in Children With Dyslexia. *Am J Speech Lang Pathol. 28*(3), 1127-1138. https://doi.org/10.1044/2019_AJSLP-18-0106.

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

Gardini, S., Cuetos, F., Fasano, F., Pellegrini, F. F., Marchi, M., Venneri, A., et al. (2013). Brain structural substrates of semantic memory decline inmild cognitive impairment. *Curr. Alzheimer Res. 10*, 373–389. https://doi.org/10.2174/1567205011310040004

Gollan, T. H., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T. L. (2007). The bilingual effect on Boston Naming Test performance. *Journal of the International Neuropsychological Society : JINS*, *13*(2), 197–208. https://doi.org/10.1017/S1355617707070038

Graves, R. E., Bezeau, S. C., Fogarty, J., & Blair, R. (2004). Boston Naming Test Short Forms: A Comparison of Previous Forms with New Item Response Theory Based Forms. *Journal of Clinical and Experimental Neuropsychology*, *26*(7), 891-902. https://doi.org/10.1080/13803390490510716

Henry, J. D., and Crawford, J. R. (2004b). Verbal fluency deficits in Parkinson's disease: a meta-analysis. *J. Int. Neuropsychol. Soc. 10*, 608–622. https://doi.org/10.1017/S1355617704104141

Hobson, V. L., Hall, J. R., Harvey, M., Cullum, C. M., Lacritz, L., Massman, P. J., Waring, S. C., O'Bryant, S. E. (2011). An examination of the Boston Naming Test: calculation of "estimated" 60-item score from 30- and 15-item scores in a cognitively impaired population. *Int J Geriatr Psychiatry*, *26*(4), 351-355. https://doi.org/10.1002/gps.2533.

Kaplan, E., Goodglass, H. and Weintraub, S. (1983). The Boston Naming Test. Philadelphia: Ed Lea and Febiger.

Kaplan, E., Goodglass, H., and Weintraub, S. (2005). Test de Vocabulario de Boston, 2 Edición Edn. Madrid: Editorial Médica Paramericana, S.A.

Kavé, G., Heled, E., Vakil, E., and Agranov, E. (2011). Which verbal fluency measure is most useful in demonstrating executive deficits after traumatic brain injury? *J. Clin. Exp. Neuropsychol. 33*, 358–365. https://doi.org/10.1080/13803395.2010.518703

Kim, S. H., Strutt, A. M., Olabarrieta-Landa, L., Lequerica, A. H., Rivera, D., De Los Reyes Aragon, C. J., Utria, O., & Arango-Lasprilla, J. C. (2018). Item analysis of the Spanish version of the Boston Naming Test with a Spanish speaking adult population from

Colombia. *Clin Neuropsychol*, *32*(sup1), 29-45. https://doi.org/10.1080/13854046.2018.1441908

Kim, H., Kim, J., Kim, D. Y., and Heo, J. (2011). Differentiating between aphasic and nonaphasic stroke patients using semantic verbal fluency measures with administration time of 30 seconds. *Eur. Neurol. 65*, 113–117. https://doi.org/10.1159/000324036

Kim, S. H., Strutt, A. M., Olabarrieta-Landa, L., Lequerica, A. H., Rivera, D., De Los Reyes Aragon, C. J., Utria, O., & Arango-Lasprilla, J. C. (2018). Item analysis of the Spanish version of the Boston Naming Test with a Spanish speaking adult population from Colombia. *The Clinical neuropsychologist*, *32*(sup1), 29–45. https://doi.org/10.1080/13854046.2018.1441908

King, D., Wittenberg, R., Patel, A., Quayyum, Z., Berdunov, V., Knapp, M. (2020). The future incidence, prevalence and costs of stroke in the UK. *Age Ageing, 49*(2), 277-282. https://doi.org/10.1093/ageing/afz163.

Kristinsson, S., Zhang, W., Rorden, C., Newman-Norlund, R., Basilakos, A., Bonilha, L., Yourganov, G., Xiao, F., Hillis, A., & Fridriksson, J. (2020). Machine learning-based multimodal prediction of language outcomes in chronic aphasia. *Hum Brain Mapp, 42*(6), 1682-1698. https://doi.org/10.1002/hbm.25321.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16,* 606–613. http://dx.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 Edition Edn. New York: McGraw-Hill Irwin.

Leite, K. S., Miotto, E. C., Nitrini, R., Yassuda, M. S. (2017). Boston Naming Test (BNT) original, Brazilian adapted version and short forms: normative data for illiterate and low-educated older adults. *Int Psychogeriatr*. *29*(5), 825-833. https://doi.org/10.1017/S1041610216001952.

Lo, J., Chan, L., Flynn, S. (2021). A Systematic Review of the Incidence, Prevalence, Costs, and Activity and Work Limitations of Amputation, Osteoarthritis, Rheumatoid Arthritis, Back Pain, Multiple Sclerosis, Spinal Cord Injury, Stroke, and Traumatic Brain Injury in the United States: A 2019 Update. *Archives of Physical Medicine and Rehabilitation*, *102*(1), 115-131. https://doi.org/10.1016/j.apmr.2020.04.001

Mahoney, F. I., & Barthel, D. W. (1965). Functional evaluation: The Barthel Index. *Maryland State Medical Journal*, *14*, 61–65.

Mitrushina, M., Boone, K. B., Razani, J., & D'Elia, L. F. (2005). Handbook of Normative Data for Neuropsychological Assessment. Oxford University Press. https://play.google.com/store/books/details?id=wNuk5A4kglgC

Mueller, K. D., Hermann, B., Mecollari, J., & Turkstra, L. S. (2018). Connected speech and language in mild cognitive impairment and Alzheimer's disease: Areview of picture description tasks. *Journal of Clinical and Experimental Neuropsychology*, *40*(9), 917-939, https://doi.org/10.1080/13803395.2018.1446513

Nebreda, M. C., García-Caballero, A., Asensio, E., Revilla, P., Rodriguez-Girondo, M., Mateos, R. (2011). A short-form version of the Boston Naming Test for language screening in dementia in a bilingual rural community in Galicia (Spain). *International Psychogeriatrics*, *23*(3), 435-41. https://doi.org/10.1017/S1041610210001481.

Neils, J., Baris, J. M., Carter, C., Dell'aira, A. L., Nordloh, S. J., Weiler, E., & Weisiger, B. (1995). Effects of age, education, and living environment on Boston Naming Test performance. *Journal of Speech, Language, and Hearing Research*, *38*(5), 1143-1149.

Olabarrieta-Landa, L., Rivera, D., Morlett-Paredes, A., Jaimes-Bautista, A., Garza, M. T., Galarza-del-Angel, J., Rodríguez, W., Rábago, B., Schebela, S., Perrin, P. B., Luna, M., Longoni, M., Ocampo-Barba, N., Aliaga, A., Saracho, C. P., Bringas, M. L., Esenarro, L., García-Egan, P., & Arango-Lasprilla, J. C. (2015). Standard form of the Boston Naming Test: Normative data for the Latin American Spanish speaking adult population. *NeuroRehabilitation*, *37*(4), 501-513. http://dx.doi.org/10.3233/NRE-151278

Ostrosky-Solís, F., López-Arango, G., & Ardila, A. (2000). Sensitivity and specificity of the Mini-Mental State Examination in a Spanish-speaking population. *Applied Neuropsychology*, *7*(1), 25-31.

Peña-Casanova, J., Quiñones-Ubeda, S., Gramunt-Fombuena, N., Aguilar, M., Casas, L., Molinuevo, J. L., Robles, A., Rodríguez, D., Barquero, M. S., Antúnez, C., Martínez-Parra, C., Frank-García, A., Fernández, M., Molano, A., Alfonso, V., Sol, J. M., & Blesa, R. (2009). NEURONORMA Study Team. Spanish Multicenter Normative Studies (NEURONORMA Project): norms for Boston naming test and token test. *Arch Clin Neuropsychol.*, *24*(4), 343-54. https://doi.org/10.1093/arclin/acp039

Rami, L., Serradell, M., Bosch, B., Caprile, C., Sekler, A., Villar, A., Canal, R., & Molinuevo, J. L. (2008). Normative data for the Boston Naming Test and the Pyramids and Palm Trees Test in the elderly Spanish population. *J Clin Exp Neuropsychol.*, *30*(1), 1-6. https://doi.org/10.1080/13803390701743954

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

Rivera, D., Olabarrieta-Landa, L., Van der Elst, W., González I., Ferrer-Cascales, R., Peñalver Guia, A. I, Rodriguez-Lorenzana, A., Galarza-Del-Angel, J., Irías Escher M.

J., & Arango-Lasprilla, J. C. (2020). Regression-Based Normative Data for Children From Latin America: Phonological Verbal Fluency Letters M, R, and P. *Assessment, 28*(1):264-276. https://doi.org/10.1177/1073191119897122

Rivera, D., Usuga, D. R., Mendoza, E. M. F., Arelis, A. A., Barajas, B. V. R., Islas, M. Á. M., ... & Arango-Lasprilla, J. C. (2022). Validation of the Norma Latina neuropsychological assessment battery in individuals with multiple sclerosis in Mexico. *Multiple Sclerosis and Related Disorders, 59*. https://doi.org/10.1016/j.msard.2022.103685

Roth, A. K., Denney, D. R., Burns, J. M., Lynch, S. G. (2018). Cognition in older patients with multiple sclerosis compared to patients with amnestic mild cognitive impairment and healthy older adults. *Neuropsychology, 32*(6), 654-663. https://doi.org/10.1037/neu0000453

Sachs, A., Rising, K., & Beeson, P. M. (2020). A Retrospective Study of Long-Term Improvement on the Boston Naming Test. *American journal of speech-language pathology, 29*(1S), 425–436. https://doi.org/10.1044/2019_AJSLP-CAC48-18-0224

Saxton, J., Ratcliff, G., Munro, C. A., Coffey, E. C., Becker, J. T., Fried, L., & Kuller, J. (2000). Normative Data on the Boston Naming Test and Two Equivalent 30-Item Short Forms. *The Clinical Neuropsychologist, 14*(4), 526-534. https://doi.org/10.1076/clin.14.4.526.7204

Tabet, S., LeBlanc, J., Frenette, L. C., Seresova, A., Laberge-Poirier, A., Alturki, A. Y., Marcoux, J., Maleki, M., de Guise, E. (2020). Early reading comprehension and speed of reading impairments in individuals with uncomplicated and complicated mild traumatic brain injury. *J Commun Disord. 88*. https://doi.org10.1016/j.jcomdis

Tamaš, O., Kostić, M., Kačar, A., Stefanova, E., Đokić, B. S., Stanisavljević, D., Milovanović, A., Đorđević, M., Glumbić, N., & Dragašević-Mišković, N. (2021).

Social Cognition in Patients With Cerebellar Neurodegenerative Disorders. *Frontiers in systems neuroscience, 15*, 664223. https://doi.org/10.3389/fnsys.2021.664223

Van Breukelen, G. J., & Vlaeyen, J. W. (2005). Norming clinical questionnaires with multiple regression: The Pain Cognition List. *Psycho- logical Assessment, 17,* 336–344. http://dx.doi.org/10.1037/1040-3590 .17.3.336

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,* 80 – 89. http://dx.doi.org/10.1017/S13556177 06060115

Vicente, S. G., Benito-Sánchez, I., Barbosa, F., Gaspar, N., Dores, A. R., Rivera, D., & Arango-Lasprilla, J. C. (2021). Normative data for Verbal Fluency and Object Naming Tests in a sample of European Portuguese adult population. *Applied Neuropsychology: Adult*. https://doi.org/10.1080/23279095.2020.1868472

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: International Journal of Methodology, 44,* 1105–1112. http://dx.doi.org/10.1007/s11135-009-9263-6

Zhang, Q., Feng, P., Weng, Y., Lu, X., Sun, Y., & Zhang, L. (2022). Development and psychometric testing of burn inpatient nursing dependency assessment scale. *Journal of advanced nursing*, *78*(10), 3483–3494. https://doi.org/10.1111/jan.15418

Zgaljardic, D. J., Oden, K. E., Dickson, S., Plenger, P. M., Lambert, M. E., & Miller, R. (2013). Naming Test of the Neuropsychological Assessment Battery: Reliability and Validity in a Sample of Patients with Acquired Brain Injury. *Archives of Clinical Neuropsychology, 28*(8), 859–865, https://doi.org/10.1093/arclin/act037

Table 1.
Sociodemographic characteristics for each country

Country	*n* Total	Age			Education			Sex	
								Woman	Man
		Mean (*SD*)	Min.	Max.	Mean (*SD*)	Min.	Max.	*n* (%)	*n* (%)
Argentina	316	45.7 (19.5)	18.0	89.0	13.8 (4.53)	5.0	24.0	222 (70)	94 (30)
Chile	245	57.4 (19.7)	18.0	90.0	9.55 (5.49)	2.0	24.0	143 (58)	102 (42)
Cuba	292	52.7 (19.4)	18.0	90.0	11.7 (3.64)	2.0	22.0	155 (73)	137 (27)
El Salvador	244	56.0 (22.1)	18.0	94.0	8.9 (5.3)	1.0	25.0	148 (61)	96 (39)
Honduras	181	48.8 (18.9)	18.0	89.0	8.6 (5.6)	1.0	24.0	114 (63)	67 (37)
Mexico	1129	53.5 (20.2)	18.0	94.0	9.3 (4.9)	1	26.0	753 (67)	375 (33)
Paraguay	258	53.2 (14.8)	18.0	79.0	9.5 (4.4)	3.0	23.0	160 (62)	98 (38)
Puerto Rico	163	49.0 (16.9)	21.0	89.0	13.4 (4.1)	2.0	25.0	90 (55)	61 (45)

Note. *n* indicates sample size. *SD* means Standard Deviation. Min. and Max. indicate minimum and maximum, respectively.

Table 2.
Spearman correlation and coefficient of determination between raw score and age, education and sex for each country.

Countries	**Age**		**Education**		**Sex**	
	r_s	r^2	r_s	r^2	r_s	r^2
Argentina	.072	.01	.46***	.21	.16*	.02
Chile	-.32***	.10	.48***	.23	.19*	.03
Cuba	-.23***	.06	.26**	.05	.27***	.07
El Salvador	-.19*	.04	.60***	.37	.13	.02
Honduras	-.12	.01	.69***	.47	.19	.04
Mexico	-.35***	.12	.45***	.20	.16***	.03
Paraguay	-.40***	.16	.55***	.30	.01	.00
Puerto Rico	-.24*	.06	.43***	.19	.07	.00

Note. r_s = Spearman's correlation between total scores for each country and age, education and sex. r^2 = coefficient of determination. * *p*<.01; ** *p* < .001; *** *p*<.0001.

Table 3.

Final multiple linear regression models and residuals values for the short-BNT total score by country.

Country	**Predictor**	***b***	***b* 95% CI [LL, UL]**	**Fit**	**Residuals**	
					$\hat{Y}_i$	***SD*e**
Argentina	(Intercept)	12.01**	[11.81, 12.22]	R^2 = .268** 95% CI[.18,.34]	*<11.5*	*0.975*
	Age	0.01**	[0.01, 0.02]		*11.6-12.1*	*1.09*
	Education	0.21**	[0.17, 0.25]		*12.2-12.9*	*0.868*
	Sex	0.66**	[0.28, 1.05]		*>12.9*	*0.728*
Chile	(Intercept)	10.76**	[10.36, 11.15]	R^2 = .293** 95% CI[.20,.37]	*<9.9*	*0.468*
	Age	-0.03**	[-0.04, -0.01]		*10.0-10.9*	*0.858*
	Education	0.21**	[0.15, 0.27]		*11.0-12.1*	*0.802*
	Sex	0.86**	[0.25, 1.47]		*>12.1*	*0.863*
Cuba	(Intercept)	11.36**	[10.98, 11.74]	R^2 = .229** 95% CI[.14,.30]	All values	*2.193*
	Education	0.22**	[0.15, 0.29]			
	Education^2	-0.03**	[-0.05, -0.02]			
	Sex	1.35**	[0.84, 1.86]			
El Salvador	(Intercept)	10.37**	[10.10, 10.63]	R^2 = .378** 95% CI[.29,.46]	All values	*2.096*
	Education	0.31**	[0.26, 0.36]			
Honduras	(Intercept)	8.66**	[8.35, 8.98]	R^2 = .488** 95% CI[.39,.57]	*<6.9*	*2.06*
	Education	0.37**	[0.31, 0.43]		*7.0-7.7*	*2.42*
					7.8-9.9	*1.90*
					>9.9	*1.88*
Mexico	(Intercept)	10.50**	[10.36, 10.65]	R^2 = .281** 95% CI[.24,.32]	All values	*2.028*
	Age	-0.03**	[-0.04, -0.03]			
	Education	0.19**	[0.16, 0.21]			
	Sex	0.50**	[0.25, 0.76]			
Paraguay	(Intercept)	13.27**	[13.06, 13.49]	R^2 = .295** 95% CI[.20,.37]	All values	*1.207*
	Age	-0.02**	[-0.03, -0.00]			
	Age^2	-1.1^{-3}**	[< -0.00, < -0.00]			
	Education	0.15**	[0.11, 0.19]			
Puerto Rico	(Intercept)	11.50**	[11.12, 11.87]	R^2 = .202** 95% CI[.09,.30]	*<11.1*	*2.79*
	Age	-0.02	[-0.05, 0.00]		*11.2-11.4*	*2.83*
	Education	0.24**	[0.14, 0.33]		*11.5-12.2*	*1.87*
					>12.3	*1.71*

Note. b represents unstandardized regression weights. *Lower Limit (LL)* and *Upper Limit* (*UL)* indicate the lower and upper limits of a confidence interval, respectively. ** $p < .001$