Journal of the International Neuropsychological Society (2026), **32**, 406–417
doi:10.1017/S1355617726101994

Research Article

Regression-based normative data for neuropsychological assessment in Buenos Aires, Argentina: RAVLT, Trail Making Test, digit span, and verbal fluency

Nicolás Corvalán[1], Florentina Morello-García[2], Samanta Leiva[3], Ricardo F. Allegri[1,2] and Lucía Crivelli[1,2]

[1]Neurology, Neuropsychology and Neuropsychiatric Department, FLENI, Argentina, [2]Institute of Neurosciences (INEU), Fleni-CONICET, Argentina and [3]Unidad de Neuropsicología HIGA Eva Perón, LICH-UNSAM, Argentina

Abstract

Objective: To develop regression-based normative data for a set of widely used neuropsychological tests and to provide a freely accessible normative data calculator for use in Buenos Aires, Argentina. **Methods:** Participants were drawn from a large clinical dataset and selected to reflect preserved global cognition, intact memory performance, and functional independence. Regression-based norms were developed for the Rey Auditory Verbal Learning Test, Verbal Fluency, Trail Making Test, and Digit Span tasks. For each model, predictors included age, sex, and years of education, as well as polynomial terms and interaction effects; final models were obtained via stepwise backward elimination procedures. Model assumptions were evaluated, and predictive performance was estimated using 10-fold cross-validation. **Results:** 393 participants met inclusion criteria. Age was negatively associated with verbal memory, fluency, and executive function, whereas higher educational attainment was associated with better performance. Women outperformed men in verbal memory and phonological fluency. Demographic variables explained between 2% and 23% of variance across tests. An open-access calculator was developed to facilitate individualized normative estimation in clinical settings. **Conclusions:** This study provides locally developed, regression-based norms for neuropsychological assessment in Buenos Aires – Argentina, addressing the lack of culturally relevant norms and limitations of traditional methods.

Keywords: Aging; executive function; neuropsychological norms; regression-based norms; verbal fluency; verbal memory

(Received 24 December 2025; final revision 3 April 2026; accepted 9 April 2026; First Published online 11 May 2026)

Statement of Research Significance

Research Question(s) or Topic(s): This study developed regression-based normative data for widely used neuropsychological tests in Argentine adults aged 50 years and older, examining the effects of age, sex, and education on cognitive performance. **Main Findings:** Age, sex, and education were significant predictors of performance across several cognitive measures. Older age was generally associated with lower performance, higher education with better performance, and female sex with better scores in verbal memory and phonological fluency. The results support regression-based models as a useful framework for adjusting neuropsychological test scores for demographic influences. **Study Contributions:** This study provides locally derived normative data for Argentina, helping reduce reliance on foreign norms in clinical practice. It also shows the methodological value of regression-based norming and offers an open-access calculator to facilitate application in clinical and research settings.

Introduction

Argentina, the second-largest country in Latin America and home to nearly 9% of the world's Latino population, is confronting the dual pressures of rapid population aging and an increasing prevalence of dementia risk factors. According to the National Institute of Statistics and Censuses (INDEC, 2024), adults aged 60 years and older now account for 16.2% of the total population, underscoring a marked demographic transition over the past century. Using a decade of data from the National Disability Registry, Calandri et al. (2025) estimated age-standardized dementia incidence rates per 100,000 inhabitants based on population projections from the 2010 and 2022 national censuses. Their analyses indicate that in 2023 Argentina registered an age-standardized incidence of 25.3 new dementia cases per 100,000 inhabitants, with higher rates among women (31.1) than men (19.1). In this national context, neuropsychological assessment

Corresponding author: Nicolás Corvalán; Email: ncorvalan@fleni.org.ar
Nicolás Corvalán and Florentina Morello-García join as first authors.

Cite this article: Corvalán N., Morello-García F., Leiva S., Allegri R.F., & Crivelli L. (2026) Regression-based normative data for neuropsychological assessment in Buenos Aires, Argentina: RAVLT, Trail Making Test, digit span, and verbal fluency. *Journal of the International Neuropsychological Society*, **32**: 406–417, https://doi.org/10.1017/S1355617726101994

© The Author(s), 2026. Published by Cambridge University Press on behalf of International Neuropsychological Society

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 20:33:32, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1355617726101994

constitutes an indispensable tool for diagnosis, treatment planning, and clinical decision-making (Soer et al., 2009). Consequently, the use of culturally adapted neuropsychological measures and locally derived normative data represents a necessary condition for valid clinical inferences.

Neuropsychological inferences and sociodemographics

Yet cognition is, by definition, demographic-dependent: its expression and measurement are inseparable from the individual's social, educational, and biological context (Fittipaldi et al., 2023). Studies with large samples consistently indicate that sociodemographic predictors explain a substantial proportion of cognitive variability across different domains, such as attention, cognitive flexibility, reasoning, processing speed, and working memory (Marzuki et al., 2024; Nichols et al., 2021; Rexroth et al., 2013). Taken together, these findings indicate that younger age and higher educational level are associated with better cognitive performance, whereas sex may be differentially related to specific domains. However, the tendency to consider the effect of sociodemographic variables on aging (Jawinski et al., 2022; Yan et al., 2015) and neurodegeneration (Pinaya et al., 2021; Verdi et al., 2022) is not a common practice. Nevertheless, sociodemographic variables such as age (Bethlehem et al., 2022; Cesnaite et al., 2023; Di Biase et al., 2023; Gaal et al., 2010; Ishii et al., 2017 ; Murty et al., 2020; Rossini et al., 2007; Stacey et al., 2021; Stier et al., 2023; Trammell et al., 2017; Trondle et al., 2023; Valdes-Sosa et al., 2021; Voytek et al., 2015; Zappasodi et al., 2015) and education (Gordon et al., 2008; Jokinen et al., 2016; Habes et al., 2021; Franzmeier et al., 2017a, 2017b) are important predictors of changes over time in the brain and behavior.

In this regard, data derived from cognitive tests must be adjusted for key sociodemographic variables such as age, sex, and education (Innocenti et al., 2023). Such adjustment enhances the representativeness of the test scores and improves the validity of test interpretation.

Normative data in Argentina

Despite their critical role, normative data (ND) – defined as reference values that allow clinicians to interpret an individual's test performance by comparison with a reference population – specifically developed for Latin American (LA) populations remain scarce and often lack sufficient methodological rigor (Guàrdia-Olmos et al., 2015; Perez-Lao et al., 2025). A national survey of neuropsychologists in Argentina ($n = 107$) reported that 58% of the professionals use ND from foreign countries, and 64% of respondents expressed concern about the absence of culturally appropriate ND (Fernández et al., 2016). In the absence of locally validated norms, clinicians frequently rely on ND developed in the United States or Europe, which may not adequately account for cultural, linguistic, and educational differences (Casaletto & Heaton, 2017; Daga et al., 2025; Karstens et al., 2024). Applying non-local norms can result in substantial diagnostic inaccuracies, increasing the risk of false positives and misdiagnoses when assessing individuals from diverse cultural backgrounds (Norman et al., 2000, 2011). For example, in a study by Daugherty et al. (2017), the application of North American norms to Colombian, Moroccan, and Spanish samples led to false positive rates of 16%, 23%, and 6%, respectively. Depending on the specific test and population, up to 51% of individuals were misclassified as having mild or major cognitive disorder. Among Colombians, 23% met criteria for Mild Cognitive Disorder and 11% for major cognitive disorder; for Moroccans, these rates were 20% and 24%, respectively, compared to only 11% and 1.2% among Spaniards.

Despite these challenges, normative research in neuropsychology in Argentina has progressed over the past two decades (Martino et al., 2021); however, the development of ND has been heterogeneous, with studies varying widely in their methodology. As an example of task-specific normative efforts, Margulis et al. (2019) developed ND for the Trail Making Test (TMT) in 407 individuals aged 16 to 90 years from Buenos Aires, providing norms stratified by age and education. More recently, Filippetti and Gutiérrez (2025), produced regression-based norms for 413 children and adolescents aged 5 to 15.

For verbal fluency tasks, several Argentine studies have been done. Butman et al. (2000) generated the first local ND for phonemic and semantic fluency in 266 adults aged 16 to 83 years, stratified by 10-year age bands and three educational levels. Marino and Alderete (2010) expanded the scope by providing norms for categorical, phonemic, grammatical, and combined fluency tasks in 259 adults aged to 15 to 70 years, stratifying by sex, age, and education groups. Martino et al. (2022) developed norms for phonemic, semantic, and action fluency in 303 adults aged 50 years and older. Although they modeled sociodemographic effects, the final normative tables were stratified by age and educational level.

With respect to memory assessment, Leis et al. (2018) published Argentine ND for the verbal subtests of the Signoret Memory Efficiency Battery (BEM 144) in 227 cognitively healthy adults and 95 patients with mild cognitive impairment (MCI). Norms were provided for six age groups. Complementarily, Román et al. (2016) validated the Argentine version of the Memory Binding Test (MBT) for early MCI detection in 88 participants (46 healthy controls and 42 amnestic MCI). The authors provided stratified ND by age (≤65 vs. ≥66 years) and education (≤12 vs. ≥13 years).

Stratified normative data vs the regression-based approach

The traditional approach to constructing neuropsychological norms involves defining a reference population and categorizing continuous variables, such as age, into discrete intervals (e.g., 60–69 years), with specific norms calculated for each. The approach involves the calculation of the sample mean and its standard deviation.

However, this approach presents limitations, as it can lead to abrupt jumps in normative scores when transitioning between age intervals (Kiselica et al., 2024). This means that nearly identical raw scores may be interpreted differently for individuals whose ages fall just within successive intervals, even if they differ by only a few days (e.g., 59 years and 364 days *vs.* 60 years and 1 day). Such jumps are unrealistic, as they contradict the theoretical understanding that score distributions change smoothly and continuously (Van Breukelen & Vlaeyen, 2005; Zachary & Gorsuch, 1985). Traditional norming relies on age intervals, which implies the implicit assumption that the test score distribution remains constant within each interval. However, when score distribution changes continuously, this assumption only holds to an adequate degree if very narrow intervals are used (Figure 1).

In contrast, regression-based normative methods use covariates – such as age, sex, and education – as independent variables to predict raw test scores. These predicted scores are then standardized to generate an empirical distribution that adjusts for the covariates of interest (Shirk et al., 2011). Unlike traditional normative methods, which often rely on arbitrary grouping,

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 20:33:32, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1355617726101994

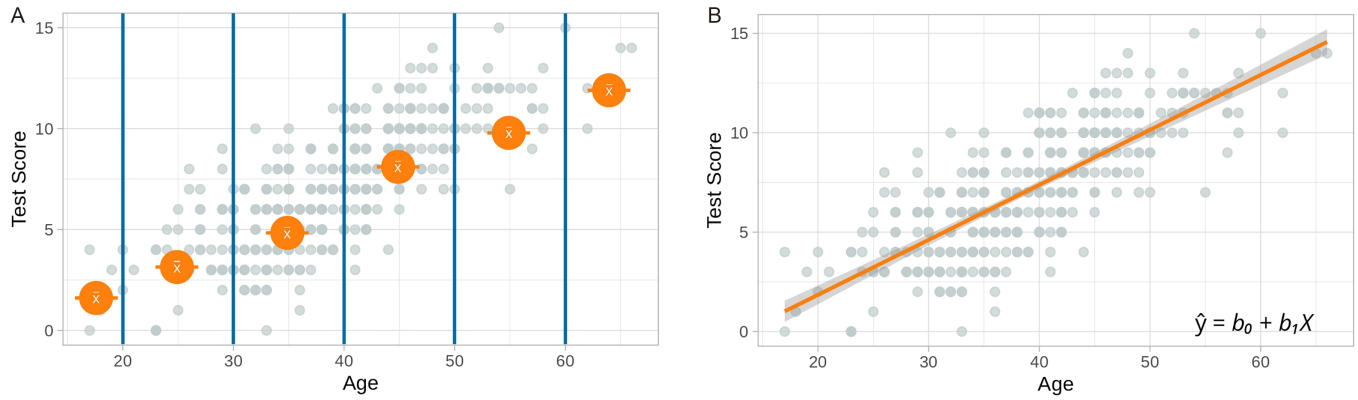


Figure 1. Panel A: Traditional normative approach using discrete age bins, where test scores are compared to group means within each age interval. **Panel B:** Regression-based normative approach, where test scores are predicted as a continuous function of covariates using linear models. Inspired by Timmerman et al., 2021.

regression-based approaches treat continuous variables as such, avoiding artificial cutoffs and better capturing individual variability. These models can incorporate multiple covariates simultaneously, be tailored to specific neuropsychological tests and populations (Timmerman et al., 2021), and include interaction effects between variables. Additionally, non-linear relationships between predictors and outcomes can be modeled to more accurately reflect complex patterns (Shirk et al., 2011). Finally, empirical estimates suggest that traditional normed score calculation requires sample sizes two and a half to five times larger than regression-based normed score calculation (Oosterhuis et al., 2016).

Due to the methodological limitations in generating norms via stratification and the lack of local ND, it becomes necessary to develop ND that ensure robustness and generalizability. Using local ND will allow for greater precision in the cognitive mapping and strengthen clinical decision-making. In this context, linear modeling offers several advantages: it quantifies the independent contribution of each predictor to cognitive performance and enables the simultaneous inclusion of multiple predictors, thereby avoiding arbitrary stratification schemes and improving model efficiency and interpretability. With this in mind, the objective of this study is to develop regression-based ND for a set of widely used neuropsychological tests. In addition, as a secondary aim, we will provide a freely accessible calculator for use in Buenos Aires, Argentina.

Materials and methods

Participants

Participants were drawn from a large clinical dataset maintained by the Fleni Neurology, Neuropsychology, and Neuropsychiatric Department (Buenos Aires, Argentina). Patients underwent evaluations between 2020 and 2024. Inclusion criteria were as follows: (a) age over 50 years, (b) native Rioplatense Spanish speaker, (c) a Montreal Cognitive Assessment (MoCA) score ≥ 26 (González Palau & Berrios, 2018), (d) z-scores > -1.5 on both immediate and delayed recall phases of the Craft Story 21 (Sachs et al., 2020), and (e) a Functional Assessment Questionnaire (FAQ) score < 9. Patients with neurological and/or psychiatric diagnoses, as determined by clinical neurologists according to Diagnostic and Statistical Manual of Mental Disorders V (DSM-V), were excluded from the study.

This project was approved by the Fleni Research Ethics Committee (registration code 16997). The study used retrospective data extracted from an institutional database, and the requirement for informed consent was waived by the ethics committee. The research was conducted in accordance with the principles of the Declaration of Helsinki.

Cognitive measures

The criteria for including the following cognitive tests were based on two main aspects: (a) their dissemination across the region, considering both accessibility and open availability of stimuli, as detailed below and (b) their adequate psychometric properties. Neuropsychological tests are part of the Uniform Data Set Neuropsychological Battery (Morris et al., 2006; Weintraub et al., 2009). The complete list of stimuli can be viewed at https://www.naccdata.org/collect-and-submit-nacc-data. The version of the RAVLT used corresponds to Burín et al. (2003).

Rey auditory verbal learning test

The Rey Auditory Verbal Learning Test (RAVLT; Lezak, 1983; Rey, 1964) assesses episodic verbal memory through the learning and recall of a word list. The test begins with a learning phase consisting of five trials during which the subject is presented with 15 words that must be recalled immediately after being pronounced by the examiner. Subsequently, a unique alternative list of 15 words (interference list) is introduced, which the subject must also recall immediately. Next, the subject is asked to recall the first list again. After a set interval (20–30 minutes), long-term retention is assessed through delayed recall, and finally, a recognition task is performed to identify the previously presented words. In addition to the scores obtained during the initial phase, which allow for the construction of a learning curve, and the results of the various recall and recognition trials necessary to evaluate the processes of encoding, storage, and retrieval, this test provides valuable information about the presence of intrusions and false positives, key aspects for constructing the subject's cognitive profile. According to Alviarez-Schulze et al. (2022), the RAVLT demonstrates adequate construct validity, with a four-factor structure showing good model fit (GFI = 0.95, RMSEA < 0.09), and evidence of convergent validity, as indicated by significant correlations between RAVLT scores and S-FNAME performance ($r = 0.37$–0.43; $p < .01$).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 20:33:32, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1355617726101994

Semantic and phonological fluency

Verbal Fluency tasks assess the subject's ability to access the lexicon and retrieve information related to words. This information may be linked to meaning (semantic fluency) or formal structure (phonological fluency). In both tasks, the subject is asked to produce as many words as possible within a set time period (1 minute), under a specific restriction. In semantic fluency, responses must be limited to a specific category; for this study, the category of animals was used (Morris et al., 1989). In phonological fluency, the subject must generate words starting with a specific letter; in this case, the letter P was chosen, as it is widely used in the region (Butman et al., 2000). These tasks not only assess language-related aspects but also involve other cognitive domains, such as executive functions. Additionally, they allow for the analysis of other elements, beyond the final score, such as the presence of perseverations, intrusions, and the order in which words are produced, providing insight into the strategies used by the subject. In an Argentine study by Martino et al. (2022), test–retest reliability was examined in a sample of 85 non-clinical adults (mean age = 73.7 years; 75.3% women). Phonological and semantic fluency tasks demonstrated good reliability, with intraclass correlation coefficients (ICCs) of 0.77 and 0.79, respectively, while the action fluency task showed excellent reliability (ICC = 0.90). Similarly, in the normative study by Butman et al. (2000), phonological and semantic fluency tasks provided evidence of convergent validity, as both measures correlated positively with years of education and total scores on the Mini-Mental State Examination (MMSE) and negatively with age.

Trail Making Test

The Trail Making Test (TMT; Reitan, 1958) is a neuropsychological assessment consisting of two parts. In Part A, sustained attention and visuospatial skills, such as visual search, are assessed by asking the subject to connect numbers from 1 to 25 in ascending order with a straight line as quickly as possible. Part B shares a similar format with Part A but is more demanding. In this task, the subject must alternate between numbers (1 to 13) and letters (A to L), connecting them with a straight line as quickly as possible while maintaining ascending numerical and alphabetical order (e.g., 1-A-2-B). In this part, alternating attention is evaluated in addition to the skills assessed in Part A. In both tasks, scoring is based on the total time taken (in seconds). Additionally, the number of errors and, particularly in Part B, the type of error may also be recorded. Regarding its psychometric properties, in an Argentine study, Fernández et al. (2002) standardized the TMT and examined its convergent validity, finding a significant association between the TMT and verbal fluency (TMT-A and verbal fluency: $r = -0.36$, $p < .001$, TMT-B and verbal fluency: $r = -0.40$, $p < .001$) and the Porteus Maze Test (TMT-A and Porteus Maze Test: $r = -0.48$, $p < .001$, TMT-B and Porteus Maze Test: $r = -0.44$, $p < .001$).

Digit span test

The Digit Span test (Richardson, 2007) is a tool frequently used in neuropsychological assessment. It consists of two parts: Digit Forward and Digit Backward. In the Digit Forward task, the subject must immediately repeat, in the same order, a series of digits that progressively increase in length as the test advances. The series begins with 2 digits and increases in number with each level. This task primarily assesses attentional span. The Digit Backward task follows a similar format in terms of presentation. The subject listens to a series of digits that also increase in length but, in this case, must repeat them in reverse order. This task assesses working memory, as it requires temporarily storing the digits and mentally manipulating them before reproducing them. Separate scores are obtained for each part, allowing for independent analysis of the assessed domains. Barreyro et al. (2015) examined the reliability and convergent validity of this test in an Argentine sample. The results indicated good internal consistency ($\alpha = 0.80$ and $\alpha = 0.83$ for the two subtests) and significant associations between the Digit Span and the Letter–Number Sequencing task. A confirmatory factor analysis supported a model with two first-order factors (one for each task) and a second-order working memory factor, showing good fit indices ($\chi^2(1) = 0.62$, $p = .43$; AGFI = 0.97; CFI = 0.99; TLI = 0.99; RMSEA $\leq$ 0.00).

Sociodemographic variables

The sociodemographic variables to be included are age, sex, and educational attainment. Age will be defined as chronological age; sex will be classified as female or male (female = 1, male = 0); and educational attainment will be expressed as the total number of completed years of formal education.

Data analysis

All analyses were conducted using R version 4.5.0. The following packages were used: tidyverse, NormData, caret, MASS, arsenal, flexplot, and forecast. The ND development procedure is based on the one proposed by Van der Elst (2024).

Stage 1: Building the regression model.

The regression-based ND were calculated using age, sex, and education as predictors and raw scores as outcomes for each variable. The linear regression model can be expressed as:

$$Y' = \alpha + \beta_1 X_1 + \beta_2 X_2 + \epsilon \qquad N(0, \sigma_\epsilon^2)$$

where Y′ is the predicted score, α is the intercept, β_i is the beta coefficient associated with variable i, and X_i is the value of predictor i. The intercept (α) represents the value of the response variable when all predictors are equal to 0. At the same time, the beta coefficients (β_i) indicate the average change in the response variable for a one-unit increase in the predictor, holding all other predictors constant.

The first main objective of Stage 1 of the regression-based normative procedure is to obtain a parsimonious model for the mean structure. Age was centered (Age^c = age-$\bar{x}$ in the sample) in order to render the model intercept clinically interpretable so that it represents the expected performance of an average individual from the normative sample rather than a value at age zero. The considered model includes Age^c, Age^{2c}, Age^{3c}, $\log(\frac{age}{100})$, Sex, Educational Level, $Age^c \times$ Sex, $Age^c \times$ Educational Level, Sex × Educational Level. The model includes three main effects (Age^c, Sex, and Educational Level), higher-order polynomial terms up to the third order for the centered quantitative independent variable (Age^c), and interaction terms.

A stepwise backward elimination model-building strategy was employed. This approach, based on the Akaike Information Criterion (AIC), starts from the full theoretical model and progressively removes predictors that do not improve model

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 20:33:32, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1355617726101994

fit and has been widely used in regression-based normative frameworks (e.g., Casaletto & Heaton, 2017; Mitrushina et al., 2005; Van der Elst et al., 2013). Following model selection, statistical significance of retained predictors was evaluated using Bonferroni-corrected p-values to control for multiple comparisons (Arcara et al., 2024).

The following assumptions were evaluated for all regression models: multicollinearity (Variance Inflation Factor [VIF] < 10), homoscedasticity (Breusch–Pagan test), normality of the standardized residuals (Kolmogorov–Smirnov test and histogram plots), kurtosis, skeweness, and the existence of influential values (measured by standardized residual variance distance). When the homoscedasticity assumption was not met, heteroskedasticity-consistent (HC) standard errors were computed using the HC0 estimator (White, 1980). Finally, model comparison was complemented using information-theoretic criteria (AIC and BIC).

In order to reduce overfitting, we additionally performed 10-fold cross-validation to estimate the out-of-sample predictive performance of each model. In this procedure, the total sample ($n = 393$) was partitioned into 10 folds of approximately equal size. Each validation fold contained between 39 and 40 participants, while the corresponding training sets ranged from 353 to 354 participants. In each iteration, 9 folds were used for model training and the remaining fold for validation. This process was repeated 10 times so that each observation contributed once to validation and nine times to training. Whereas the full-sample R^2 reflects explained variance, the cross-validated R^2, RMSE, and MAE quantify the expected accuracy when predicting new subjects.

Stage 2: Deriving normative data

The procedure consists of:

1. Calculating $\widehat{Y}_0$ for the evaluated individual $i = 0$ using the Stage 1 prediction function.
2. Calculating the residuals

$$\varepsilon_0 = Y_0 - \widehat{Y}_0$$

3. Calculating the standardized residual

$$\left(\delta = \frac{\varepsilon}{\sigma_\varepsilon}\right)$$

4. The individual residual score (δ_0), obtained as described in Step 3, was converted into an estimated percentile rank (π_0) using the empirical cumulative distribution of residuals derived from the normative sample:

$$\pi_0 = \Psi(\delta_0)$$

Results

Sociodemographics

The initial dataset consisted of a total of 13,729 patients. After applying the eligibility criteria, 393 were selected. A sample size evaluation is provided in the Supplementary Materials. Sociodemographic characteristics of the sample are described in Table 1. Neuropsychological test results are described in Table 2.

Table 1. Sociodemographic data ($n = 393$) grouped by age, sex, and educational level

	50–60		61–70		71–80		>81		
Educational level	F	M	F	M	F	M	F	M	Total
Less than 12 years	2	1	16	17	39	25	15	4	**119**
More than 12 years	12	4	62	41	49	71	15	20	**274**
Total	**14**	**5**	**78**	**58**	**88**	**96**	**30**	**24**	**393**

Note: F = Female; M = Male.

Model assumptions

Across all models, multicollinearity – measured through generalized variance inflation factors – was below conventional thresholds. No models showed evidence of problematic collinearity. For the RAVLT Learning, RAVLT Delayed, Semantic Fluency, Phonological Fluency, Digit Span Forward, Trail Making Test A and Trail Making Test B models, the assumption of homoscedasticity was supported by non-significant Breusch–Pagan tests. In these models, residual normality was generally adequate, although mild deviations from normality were observed in some cases (e.g., Semantic and Phonological fluency, Digit Span Forward, and Trail Making Test B). No influential outliers were detected.

In contrast, the Digit Span Backward model showed evidence of heteroscedasticity, as indicated by Breusch–Pagan tests. For this model, heteroscedasticity-consistent standard errors were therefore used. Detailed test statistics for all assumption checks and model comparisons are reported in the Supplementary Materials section (Appendix 2–3).

Model performance and predictors

Across verbal memory, fluency, working memory, and executive measures, demographic variables accounted for a systematic proportion of performance variability. Parameter estimates, explained variance proportion (R^2 and adjusted R^2), error metrics, confidence intervals, and the final formula for all adjusted models are presented in Table 3. Interaction terms and higher-order age polynomials were not retained in the final models due to lack of statistical significance.

For verbal memory, age, sex, and years of education jointly contributed to RAVLT performance. Increasing age was associated with lower scores, whereas female sex and higher educational attainment were associated with better outcomes. The RAVLT learning model explained 23% of the variance, with ten-fold cross-validation yielding an R^2 of 0.24 (RMSE = 7.86; MAE = 6.23). For RAVLT Delayed Recall, the model explained 17% of the variance, and cross-validation yielded an R^2 of 0.17 (RMSE = 2.80; MAE = 2.28).

In verbal fluency tasks, age (modeled as log(age/100)), education and sex were the primary determinants of performance. Phonological Fluency was associated with age, sex – with females outperforming males – and years of education. The proportion of explained variance was 8%, and cross-validation results were consistent with the fitted model (R^2 = 0.08; MSE = 4.34; MAE = 3.43). Semantic Fluency scores decreased with increasing age and increased with higher years of education, with the final model explaining 12% of the variance and cross-validated performance remaining consistent (R^2 = 0.12; RMSE = 4.38; MAE = 3.53).

Years of education was a significant predictor of Digit Span Forward performance, accounting for approximately 2% of the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 20:33:32, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1355617726101994

Table 2. Neuropsychological and functional tests scores

	50–60M (SD)	61–70M (SD)	71–80M (SD)	>80M (SD)
Rey Auditory Verbal Learning Test				
I-V Learning phase	36.89 (8.78)	42.19 (4.83)	37.31 (8.47)	33.76 (8.34)
Delayed Recall phase	8.72 (3.16)	7.89 (2.74)	6.55 (3.02)	5.27 (3.06)
Verbal fluency				
Phonological Fluency	17.94 (4.74)	16.2 (4.56)	15.6 (4.35)	15.3 (4.64)
Semantic Fluency	22.09 (5.66)	20.08 (4.39)	18.7 (4.38)	16.92 (4.82)
Digit Span test				
Forward	6.09 (1.06)	6.14 (1.09)	6.15 (1)	5.65 (1.01)
Backward	4.38 (1.16)	4.59 (0.92)	4.37 (0.92)	4.16 (0.69)
Trail Making Test				
TMT – A	32.71 (12.61)	39.3 (13.46)	46.88 (17.36)	52.16 (14.66)
TMT – B	84.58 (40.07)	95.09 (39.34)	116.59 (51.44)	141.95 (55.03)
FAQ	1.09 (1.73)	1.47 (2.32)	1.27 (1.81)	2.57 (2.82)
MoCA	27.95 (1.39)	27.38 (1.25)	27.24 (1.07)	26.92 (1.06)
Craft Story 21				
Immediate Recall	18.79 (3.92)	19.63 (4.27)	19.41 (4.83)	17.41 (4.77)
Delayed Recall	17.11 (4.11)	17.27 (4.36)	16.21 (4.81)	14.62 (4.61)

Note: M = Mean, SD = Standard Deviation, TMT = Trail Making Test, FAQ = Functional Assessment Questionnaire, MoCA = Montreal Cognitive Assessment.

variance. Cross-validation showed a $R^2 = 0.04$, RMSE = 1.02, and MAE = 0.78. On the other hand, Digit Span Backward performance was associated with years of education, with an explained variance of 2% and cross-validation yielding similar results ($R^2 = 0.03$; RMSE = 0.90; MAE = 0.75).

Due to the positively skewed distribution of completion times, both Trail aking Test outcomes were log-transformed prior to modeling. Finally, both Trail Making Test were age-dependent. For TMT-A, older age was associated with longer completion times, with the final model explaining approximately 14% of the variance. Cross-validation results indicated comparable performance ($R^2 = 0.16$; RMSE = 0.33; MAE = 0.26). For TMT-B, performance reflected the effects of age and education, with completion times increasing with age and decreasing with higher educational attainment. The final model explained 16% of the variance, and cross-validation yielded an R^2 of 0.18 (RMSE = 0.36; MAE = 0.29). The table linking percentiles with their corresponding residual values is provided in the Supplementary Materials (Appendix 2).

Open access calculator

The four-step normative procedures explained above offer clinicians the ability to determine a *z*-score and the corresponding percentile for a participant who has a specific score on the RAVLT Learning, RAVLT Delayed, Verbal Fluency, TMT A-B, and Digit Span. To facilitate its use, the authors developed a calculator in Microsoft Excel where clinicians should include patient information (raw score for the specific test score, age, education, and sex). This excel file is freely available for all users and may be downloaded at https://github.com/nicorvfl/RegNorm. In addition, a brief tutorial on the use of the calculator, provided in both English and Spanish, is included in the Supplementary Materials (Appendix 1).

Discussion

The primary aim of this study was to develop regression-based ND for an adult Argentine population aged 50 to 90 years for the most commonly used neuropsychological tests in the region. Specifically, these norms are intended for individuals primarily residing in Buenos Aires, with demographic, educational, and socioeconomic characteristics broadly comparable to those represented in the present sample. In addition, we provide a freely accessible open access calculator. Using a clinical sample of 393 participants from Fleni, we established ND that account for sociodemographics, such as age, sex, and education. The selected tests consist of the RAVLT, Verbal Fluency (Semantic and Phonological), TMT A - B, and Digit Span (Forward and Backward). Overall, our results indicate that sociodemographic variables constitute relevant predictors of cognitive performance and that, consequently, adjusting ND for these variables represents a fundamental methodological step.

Sociodemographic factors have had a significant effect on all models of cognitive tests. Among them, age has been negatively associated with most tasks, indicating that older age is linked to lower cognitive performance. This result is consistent with previous studies (Deary et al., 2009; Murman, 2015; Zaninotto et al., 2018). Age is, as expected, one of the most important sociodemographic variables for statistical control in the generation of ND due to its effect on the structural and functional changes in the human brain. According to the global decline approach, age negatively affects cognitive control function – i.e., the ability to regulate and coordinate thoughts, actions, and emotions to achieve goals and adapt to changing environmental demands – through the slowing down of processing speed (Sorel & Pennequin, 2008). This deceleration affects cognitive skills such as attention, memory, and executive functions (Korotkevich et al., 2015) and particularly hinders the ability to encode new information effectively and, therefore, retrieve information from long-term memory (Salthouse, Atkinson, Berish, et al., 2003; Salthouse, 2009; Spreng & Turner, 2019). Global decline, in turn, may affect working memory and impact daily activities such as decision-making and problem-solving (Vaughan & Giovanello, 2010). In our article, we found that the effect of age has been transversal to cognitive function: it plays a role in memory, executive function, verbal fluency, and tasks related to processing speed.

Another demographic factor that has emerged as a relevant predictor is education. Its association with cognitive performance has been positive: as education increases, cognitive performance improves. Years of education are an important variable for explaining the variability in aging trajectories and cognitive function, and the influence of education on cognitive aging has been well documented (Galama et al., 2018, Opdebeeck et al., 2016; Chen et al., 2019; Clouston et al., 2012; Clouston et al., 2020;

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 20:33:32, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1355617726101994

Table 3. Fitted model results

Model fitted	β (95% CI)	Standard Error	*t* Value	*p* Value	$R^2 (R^2_{adj})$
RAVLT Learning					
*RAVLT Learning = 29.54 − 0.43 * (age −71.28) + 4.61 * (Sex) + 0.46 * (years of education)*					
Model parameters					0.23 (0.22)
(Intercept)	29.54 (25.24, 33.84)	2.18	13.50	<.001	
Age centered	−0.43 (−0.54, −0.33)	0.05	−8.14	<.001	
Sex [female]	4.61 (3.03, 6.19)	0.80	5.73	<.001	
Years of education	0.46 (0.19, 0.72)	0.13	3.37	<.001	
RAVLT Delayed					
*RAVLT Delayed = 4.70 − 0.13 * (age −71.28) + 1.40 * (Sex) + 0.10 * (years of education)*					
Model parameters					0.17 (0.16)
(Intercept)	4.70 (3.17, 6.23)	0.77	6.04	<.001	
Age	−0.13 (−0.16, −0.09)	0.01	−6.84	<.001	
Sex [female]	1.40 (0.83, 1.96)	0.28	4.89	<.001	
Years of education	0.10 (0.01, 0.20)	0.04	2.26	.02	
Phonological Fluency					
*Phonological Fluency = 9.34 − 5.41 * log(age / 100) + 1.92 * (Sex) + 0.24 * (years of education)*					
Model parameters					0.08 (0.08)
(Intercept)	9.34 (6.72, 11.95)	1.32	7.03	<.001	
log(age/100)	−5.41 (−9.48, −1.34)	2.07	−2.61	.009	
Sex [female]	1.92 (1.05, 2.79)	0.44	4.35	<.001	
Years of education	0.24 (0.10, 0.39)	0.07	3.32	<.001	
Semantic Fluency					
*Semantic Fluency = 11.19 − 13.21 * log(age/100) + 0.23* (years of education)*					
Model parameters					0.12 (0.11)
(Intercept)	11.19 (8.61, 13.77)	1.31	8.53	<.001	
log(age/100)	−13.21 (−17.34, −9.07)	2.10	−6.28	<.001	
Years of education	0.23 (0.08, 0.38)	0.07	3.14	.001	
Trail Making Test – *A*					
$\widehat{log(TMT - A)} = 4.15 + 1.31 * log(age/100)$					
Model parameters					0.14 (0.14)
(Intercept)	4.15 (4.04, 4.26)	0.05	72.44	<.001	
log(age/100)	1.31 (1.00, 1.63)	0.15	8.27	<.001	
Trail Making Test – B					
$\widehat{log(TMT - B)} = 4.89 + 0.02 * (age - 71.28) - 0.01 * (years\ of\ education)$					
Model parameters					0.16 (0.16)
(Intercept)	4.89 (4.70, 5.08)	0.09	50.03	<.001	
Age	0.02 (0.01, 0.02)	0.00	7.94	<.001	
Years of education	−0.01 (−0.03, −0.00)	0.00	−3.07	.002	
Digit Span Forward					
*Digit Span Forward = 5.18 + 0.06 * (years of education)*					
Model parameters					0.02 (0.02)
(Intercept)	5.18 (4.65, 5.72)	0.27	19.16	<.001	
Years of education	0.06 (0.02, 0.09)	0.01	3.40	<.001	
Digit Span Backward					
*Digit Span Backward = 3.64 + 0.05 * (years of education)*					
Model parameters					0.02 (0.02)
(Intercept)	3.64 (3.17, 4.11)	0.23	15.33	<.001	
Years of education	0.05 (0.02, 0.08)	0.01	3.39	.007	

Note: β coefficients are unstandardized.

Lövdén et al., 2020; Schneeweis, Skirbekk, & Winter-Ebmer, 2014; Wilson et al., 2019). People with higher educational levels tend to be, on average, healthier, live longer, and experience fewer instances of morbidity throughout their lives (Davies et al., 2018). The component still up for debate is whether education is a correlational or rather a causal factor. Furthermore, there are several suggestions that higher levels of education in life not only affect cognitive ability but also may attenuate aging-associated declines in cognition (Livingston et al., 2024), although there is doubt regarding the latter (Lövdén et al., 2020).

Finally, sex emerged as another significant predictor of cognitive performance. In our sample, the observed trend indicates that women showed higher cognitive performance in verbal memory tasks and phonological verbal fluency. In non-clinical populations, women tend to demonstrate better performance in verbal episodic memory functions (Lundervold et al., 2014; Sunderaraman et al., 2013; Herlitz & Yonker, 2002). Moreover, a meta-analysis by Hirnstein et al. (2023), including a sample of 355,173 participants, found that women outperform men in phonological fluency tasks but not in semantic fluency. Women also outperformed men in recall and recognition tasks.

In our study, we did not find a significant interactive effect between sociodemographic variables and cognitive performance. At this point, the literature remains at least inconclusive. In a cohort study conducted by Reas et al. (2017), involving a sample of 2,225 community-dwelling participants aged 31–99 years at their initial cognitive assessment, interactive effects between age and sex were identified for trajectories of the Mini-Mental State Examination screening test and the TMT-B, with the latter showing a steeper decline in women than in men. However, in the same study, no significant interactions between demographic variables were observed for other cognitive measures, such as the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 20:33:32, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1355617726101994

TMT-A, Verbal Fluency, or the Buschke Selective Reminding Test. On the other hand, Heiskanen et al. (2024) examined interactions among sex, age, and educational level in a longitudinal cohort from the Cardiovascular Risk in Young Finns Study. In that study, the sex×age and age×education interaction terms were not significant for any specific cognitive domain nor for overall cognitive performance. Nevertheless, a significant interaction between sex and educational level was identified for overall cognition, learning and memory, and information processing speed. In our sample, several explanations may be plausible. First, interaction effects tend to be domain-specific, and therefore large samples may be required to obtain stable and replicable estimates (Castillo et al., 2026). Second, differences in sampling characteristics may lead to inconsistencies across studies and difficulties in replicating interactive effects (Holzmeister et al., 2024; Olsson-Collentine et al., 2020). Taken together, and considering the heterogeneity of the existing evidence, further studies with greater robustness are likely needed to determine whether interactive effects are constrained to specific cognitive domains or are instead driven by particular population characteristics, and ultimately to clarify whether their influence on cognitive performance is additive or multiplicative in nature.

Finally, and in line with the previously stated need to adjust cognitive models for sociodemographic variables, it is important to note that several models (e.g., Digit Span) exhibit a low proportion of explained variance (2%). Given that a prior study reported a similar R^2 value for this task (Iñesta et al., 2021), it is reasonable to infer that performance may depend on factors not captured by variables such as age, sex, or education, including individual differences in attentional processes or cognitive reserve. In the same vein, the discrepancies observed between the in-sample R^2 and the cross-validated R^2 in models with low explained variance (e.g., cross-validated $R^2 = 0.04$ vs. $R^2 = 0.02$) should not be interpreted as evidence of superior generalization. Rather, they reflect the inherent instability of estimation under low-fit conditions, which is consistent with limited predictive capacity. This pattern contrasts with that observed in other models, such as the RAVLT, where the convergence between both R^2 suggests greater stability of the estimates. Accordingly, in models with low explained variance, both the estimates and the resulting normative predictions should be interpreted with caution. In this context, the findings of the present study indicate that the relevance of adjusting for sociodemographic variables could be task-dependent: while some measures show low explained variance, others exhibit more consistent associations. Therefore, and in light of this differential pattern, the development of ND that incorporate sociodemographic variables into the modeling becomes particularly relevant.

Another important issue addressed in this article concerns the assumptions of the linear regression model. One of the most relevant assumptions is homoscedasticity. When this assumption is violated, standard error estimates may be biased, leading the model to overestimate or underestimate predictive precision in specific regions of the predictor space. To address violations of the homoscedasticity assumption, heteroskedasticity-consistent estimators were used to obtain robust standard errors for the regression coefficients. Another relevant assumption of the linear regression model is the normality of the residuals. This assumption implies that, once the systematic component of the model has been accounted for, the residuals are approximately normally distributed around zero. From an inferential perspective, violations of this assumption may affect the accuracy of classical t-tests and confidence intervals in small samples. However, previous work has demonstrated that even under extreme departures from normality, linear regression remains robust in moderate to large samples (Lumley et al., 2002). Consequently, residual non-normality is not expected to substantially bias the estimation of average effects in samples of sufficient size, although it may still affect the interpretation of extreme individual values.

This study addresses several critical gaps in the local neuropsychological field. First, there is a substantial lack of locally developed ND in Latin America, particularly in Argentina (Perez-Lao et al., 2025). As a result, clinicians are often compelled to rely on norms derived from foreign populations (Fernández et al., 2016). This limitation is especially relevant given the central role of neuropsychological assessment in the detection of subtle cognitive changes that inform diagnosis and treatment planning. In this regard, the present study provides a robust estimation of cognitive performance based on locally derived ND. Moreover, the availability of a ND-calculator ensures that these data can be implemented across diverse clinical and research settings. Second, traditional approaches to ND construction – typically based on stratification by age, sex, and education – may fail to adequately capture the continuous nature of cognitive performance and can introduce artificial score discontinuities at subgroup boundaries. In contrast, regression-based normative methods offer a statistically more efficient and ecologically valid alternative by modeling raw test scores as continuous functions of demographic variables.

Finally, our study addresses a common issue in our region, namely the use of foreign ND. By way of example, a study by Arango-Lasprilla et al. (2015) developed regression-based ND for a series of Latin American countries (Argentina, Bolivia, Chile, Cuba, El Salvador, Guatemala, Honduras, Mexico, Paraguay, Peru, and Puerto Rico). That study illustrates the notable differential impact of sociodemographic variables on cognition, even among countries that share the same language. One such example is sex, for which the proportion of explained variance ranges from 1% to 23% across countries at different altitudes. It is therefore necessary to develop ND grounded in the national context and to discourage the importation of foreign norms.

This study has several limitations that should be acknowledged. A primary concern relates to the clinical nature of the sample. Although participants were selected to represent cognitively healthy individuals, recruitment through clinical sites may limit the generalizability of the regression estimates and introduce potential biases in error structure and measurement. To mitigate this issue, strict inclusion criteria were applied, including normal performance on a global cognitive screening measure (MoCA) and on a sensitive verbal memory task (Craft Story), which is frequently affected in the early stages of neurodegenerative disorders. In addition, intact functional status was required, as assessed by the FAQ, a measure with strong psychometric properties and demonstrated accuracy in discriminating normal cognition from mild cognitive impairment and dementia (González et al., 2022). In this line, it should be acknowledged that the inclusion of the Craft Story 21 criterion relies on ND that have not been developed at the national level. Future studies will need to establish ND for logical memory tests within the region. Another limitation concerns sample representativeness. Despite its relatively large size, the sample may not fully capture the regional, educational, and socioeconomic diversity of the Argentine population. Likewise, our normative sample consists predominantly of individuals older than 60 years. Caution is

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 20:33:32, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1355617726101994

therefore warranted when generalizing these findings to populations aged 50 to 59 years, and the potential estimation error of the model coefficients should be taken into account. Future studies should aim to expand recruitment across provinces and cultural contexts in order to enhance population-level validity. Finally, the use of linear models – particularly Gaussian models – in neuropsychological testing overlooks the structural assumption regarding the bounded nature of the dependent variable. Because cognitive scores often exhibit lower and upper limits, this approach may be suboptimal, especially at the extremes of the distribution. Future studies should employ alternative approaches, such as Bayesian methods, which allow for the explicit modeling of bounded dependent variables through distributions with restricted support, the incorporation of relevant information about score structure and variance, and the derivation of more robust inferences.

Future studies could address these limitations by incorporating multi-center samples and by more comprehensively evaluating the validity of regression-based ND using explicit training, testing, and validation frameworks. In addition, the use of penalized linear models, such as LASSO or Elastic Net, may be advantageous for variable selection and parameter estimation in larger and more heterogeneous datasets, particularly when the number of predictors is relatively high and multicollinearity among sociodemographic variables is present. Although model specification may become less transparent - especially when tuning parameters are selected via cross-validation – predicted scores can still be derived from the fitted models. In such cases, the availability of accessible computational tools (e.g., open-access calculators) becomes particularly important to ensure the practical implementation of the resulting ND.

Conclusion

This study could represent an important contribution to neuropsychological practice in Latin America by providing a regression-based normative framework and an accessible tool for cognitive assessment. By addressing critical gaps in local ND and implementing a more ecologically valid statistical approach, this work enhances the accuracy and applicability of neuropsychological evaluations in Buenos Aires – Argentina. Future research should consider additional cognitive domains and leveraging more advanced modeling techniques in order to improve ND representativeness.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S1355617726101994.

Acknowledgements. None.

Funding statement. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing interests. The authors declare that they have no competing interests.

References

Alviarez-Schulze, V., Cattaneo, G., Pachón-García, C., Solana-Sánchez, J., Tormos, J. M., Pascual-Leone, A., & Bartrés-Faz, D. (2022). Validation and normative data of the Spanish version of the Rey Auditory Verbal Learning Test and associated long-term forgetting measures in middle-aged adults. *Frontiers in Aging Neuroscience*, *14*, 809019.

Arango-Lasprilla, J. C., Rivera, D., Rodríguez, G., Garza, M. T., Galarza-Del-Angel, J., Rodríguez, W., Velázquez-Cardoso, J., Aguayo, A., Schebela, S., Weil, C., Longoni, M., Aliaga, A., Ocampo-Barba, N., Saracho, C. P., Panyavin, I., Esenarro, L., Martínez, C., García de la Cadena, C., & Perrin, P. B. (2015). Symbol digit modalities test: Normative data for the Latin American Spanish speaking adult population. *NeuroRehabilitation*, *37*(4), 625–638.

Arcara, G. (2024). Improving equivalent scores for clinical neuropsychology: A new method for regression model selection. *Neurological sciences: official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, *45*(12), 5685–5695.

Barreyro, J. P., Injoque-Ricle, I., González, J. M., & Burin, D. I. (2015). Estudio acerca de las propiedades psicométricas de pruebas clásicas de memoria de trabajo para tomas en grupos. *Anuario de investigaciones*, *22*(2), 283–288.

Bethlehem, R. A. I., Seidlitz, J., White, S. R., Vogel, J. W., Anderson, K. M., Adamson, C., Adler, S., Alexopoulos, G. S., Anagnostou, E., Areces-Gonzalez, A., Astle, D. E., Auyeung, B., Ayub, M., Bae, J., Ball, G., Baron-Cohen, S., Beare, R., Bedford, S. A., Benegal, V., . . . , Alexander-Bloch, A. F. (2022). Brain charts for the human lifespan. *Nature*, *604*, 525–533.

Burin, D. I., Ramenzoni, V., & Arizaga, R. L. (2003). Evaluación neuropsicológica del envejecimiento: normas según edad y nivel educacional. *Rev Neurol Argentina*, *28*, 149–152.

Butman, J., Allegri, R. F., Harris, P., & Drake, M. (2000). Spanish verbal fluency: Normative data in Argentina. *Medicina (Buenos Aires)*, *60*(5), 561–564.

Calandri, I. L., Suemoto, C. K., Allegri, R., & Equipo de Estadística e Investigación social, Agencia Nacional de Discapacidad, Argentina. (2025). A Decade-Long Analysis of Dementia Incidence in Argentina: Insights from a Public Registry Study. *Alzheimers & Dementia*, *20* (Suppl 7), e090532.

Casaletto, K. B., & Heaton, R. K.2017). Neuropsychological assessment: Past and future. *Journal of the International Neuropsychological Society: JINS*, *23*(9–10), 778–790.

Castillo, A., Miller, J. D., Vize, C., Baranger, D. A. A., & Lynam, D. R. (2026). When do interaction/moderation effects stabilize in linear regression? *Advances in Methods and Practices in Psychological Science*, *9*(1), 25152459251407860. https://doi.org/10.1177/25152459251407860

Cesnaite, E., Steinfath, P., Jamshidi Idaji, M., Stephani, T., Kumral, D., Haufe, S., Sander, C., Hensch, T., Hegerl, U., Riedel-Heller, S., Röhr, S., Schroeter, M. L., Witte, A. V., Villringer, A., & Nikulin, V. V. (2023). Alterations in rhythmic and non-rhythmic resting-state EEG activity and their link to cognition in older age. *NeuroImage*, *268*, 119810.

Chen, Y., Lv, C., Li, X., Zhang, J., Chen, K., Liu, Z., Li, H., Fan, J., Qin, T., Luo, L., & Zhang, Z. (2019). The positive impacts of early-life education on cognition, leisure activity, and brain structure in healthy aging. *Aging*, *11*(14), 4923–4942. https://doi.org/10.18632/aging.102088

Clouston, S. A., Kuh, D., Herd, P., Elliott, J., Richards, M., & Hofer, S. M. (2012). Benefits of educational attainment on adult fluid cognition: International evidence from three birth cohorts. *International Journal of Epidemiology*, *41*(6), 1729–1736, [PubMed: 23108707].

Clouston, S. A. P., Smith, D. M., Mukherjee, S., Zhang, Y., Hou, W., Link, B. G., & Richards, M. (2020). Education and cognitive decline: An integrative analysis of global longitudinal studies of cognitive aging. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *75*(7), e151–e160. https://doi.org/10.1093/geronb/gbz053

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *University of Chicago Legal Forum*, *1989*(1), 139–167.

Daga, M., Mohanty, P., Krishna, R., & Priya, S. R. M. (2025). Statistical validation in cultural adaptations of cognitive tests: A multi-regional systematic review [Preprint]. *arXiv*. https://doi.org/10.48550/arXiv.2504.13495

Daugherty, J. C., Puente, A. E., Fasfous, A. F., Hidalgo-Ruzzante, N., & Pérez-Garcia, M. (2017). Diagnostic mistakes of culturally diverse individuals when using North American neuropsychological tests. *Applied Neuropsychology: Adult*, *24*(1), 16–22.

Davies, N. M., Dickson, M., Smith, G. D., van den Berg, G. B., & Windmeijer, F. (2018). The causal effects of education on health outcomes in the UK Biobank. *Natural Human Behaviour*, *2*(2), 117–125.

Deary, I. J., Corley, J., Gow, A. J., Harris, S. E., Houlihan, L. M., Marioni, R. E., & Starr, J. M. (2009). Age-associated cognitive decline. *British Medical Bulletin*, *92*(1), 135–152.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 20:33:32, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1355617726101994

Di Biase, M. A., Tian, Y. E., Bethlehem, R. A. I., Seidlitz, J., Alexander-Bloch, A. F., Yeo, B. T. T., & Zalesky, A. (2023). Mapping human brain charts cross-sectionally and longitudinally. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(20), e2216798120. https://doi.org/10.1073/pnas.2216798120

Fernandez, A. L., Ferreres, A., Morlett-Paredes, A., Rivera, D., & Arango-Lasprilla, J. C. (2016). Past, present, and future of neuropsychology in Argentina. *The Clinical Neuropsychologist*, *30*(8), 1154–1178. https://doi.org/10.1080/13854046.2016.1197313

Fernández, A. L., Marino, J. C., & Alderete, A. M. (2002). Estandarización y validez conceptual del test del trazo en una muestra de adultos argentinos. *Revista Neurológica Argentina*, *27*(2), 83–88.

Filippetti, V. A., & Gutiérrez, M. (2025). Interpreting the direct- and derived-trail making test scores in Argentinian children: Regression-based norms, convergent validity, test–retest reliability, and practice effects. *The Clinical Neuropsychologist*, *39*(6), 1696–1721.

Fittipaldi, S., Legaz, A., Maito, M., Hernandez, H., Altschuler, F., Canziani, V., Moguilner, S., Gillan, C., Castillo, J., Lillo, P., Custodio, N., Avila-Funes, J., Cardona, J., Slachevsky, A., Henriquez, F., Fraile-Vazquez, M., de Souza, L. C., Borroni, B., Hornberger, M., . . . , Ibanez, A. (2023). Heterogeneous factors influence social cognition across diverse settings in brain health and age-related diseases. Research square, rs.3.rs-3007086.

Franzmeier, N., Duering, M., Weiner, M., Dichgans, M., & Ewers, M. (2017a). Alzheimer's disease neuroimaging I. Left frontal cortex connectivity underlies cognitive reserve in prodromal Alzheimer disease. *Neurology*, *88*(11), 1054–1061. https://doi.org/10.1212/WNL.0000000000003711

Franzmeier, N., Göttler, J., Grimmer, T., Drzezga, A., Áraque-Caballero, M. A., Simon-Vermot, L., Taylor, A. N. W., Bürger, K., Catak, C., Janowitz, D., Müller, C., Duering, M., Sorg, C., & Ewers, M. (2017b). Resting-state connectivity of the left frontal cortex to the default mode and dorsal attention network supports reserve in mild cognitive impairment. *Frontiers in Aging Neuroscience*, *9*, 264. https://doi.org/10.3389/fnagi.2017.00264

Gaál, Z. A., Boha, R., Stam, C. J., & Molnár, M. (2010). Age-dependent features of EEG-reactivity—Spectral, complexity, and network characteristics. *Neuroscience Letters*, *479*(1), 79–84.

Galama, T., Lleras-Muney, A., & Kippersluis, H. V. (2018). The effect of education on health and mortality: A review of experimental and quasi-experimental evidence. *Oxford Research Encyclopedia of Economics and Finance*.

González, D. A., Gonzales, M. M., Resch, Z. J., Sullivan, A. C., & Soble, J. R. (2022). Comprehensive evaluation of the functional activities questionnaire (FAQ) and its reliability and validity. *Assessment*, *29*(4), 748–763.

González Palau, F., Berrios, W., García Basalo, M. M., Ojea Quintana, M., Fernández, M., García Basalo, M. J., Fernández, M. C., Cámpora, N., Matusevich, L. D., & Golimstok, Á. (2018). Validez de la prueba de Evaluación Cognitiva de Montreal (MoCA) como una herramienta de screening para detectar Deterioro Cognitivo Leve en población de Buenos Aires, Argentina. *Vertex Revista Argentina De Psiquiatría*, *29*(140), 261–269. Recuperado a partir de. https://revistavertex.com.ar/ojs/index.php/vertex/article/view/300

Gordon, B. A., Rykhlevskaia, E. I., Brumback, C. R., Lee, Y., Elavsky, S., Konopack, J. F., McAuley, E., Kramer, A. F., Colcombe, S., Gratton, G., & Fabiani, M. (2008). Neuroanatomical correlates of aging, cardiopulmonary fitness level, and education. *Psychophysiology*, *45*(5), 825–838.

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.

Habes, M., Pomponio, R., Shou, H., Doshi, J., Mamourian, E., Erus, G., Nasrallah, I., Launer, L. J., Rashid, T., Bilgel, M., Fan, Y., Toledo, J. B., Yaffe, K., Sotiras, A., Srinivasan, D., Espeland, M., Masters, C., Maruff, P., Fripp, J., . . . , iSTAGING consortium, the Preclinical AD consortium, the ADNI, and the CARDIA studies (2021). The Brain Chart of Aging: Machine-learning analytics reveals links between brain aging, white matter disease, amyloid burden, and cognition in the iSTAGING consortium of 10,216 harmonized MR scans. *Alzheimer's & Dementia*, *17*(1), 89–102.

Heiskanen, M. A., Nevalainen, J., Pahkala, K., Juonala, M., Hutri, N., Kähönen, M., Jokinen, E., Laitinen, T. P., Tossavainen, P., Taittonen, L., Viikari, J. S. A., Raitakari, O. T., & Rovio, S. P. (2024). Change in cognitive performance during seven-year follow-up in midlife is associated with sex, age, and education - The Cardiovascular Risk in Young Finns Study. *Journal of Neurology*, *271*(8), 5165–5176. https://doi.org/10.1007/s00415-024-12466-2

Herlitz, A., & Yonker, J. E. (2002). Sex differences in episodic memory: The influence of intelligence. *Journal of Clinical and Experimental Neuropsychology*, *24*(1), 107–114. https://doi.org/10.1076/jcen.24.1.107.970

Hirnstein, M., Stuebs, J., Moè, A., & Hausmann, M. (2023). Sex/gender differences in verbal fluency and verbal-episodic memory: A meta-analysis. *Perspectives on Psychological Science*, *18*(1), 67–90.

Holzmeister, F., Johannesson, M., Böhm, R., Dreber, A., Huber, J., & Kirchler, M. (2024). Heterogeneity in effect size estimates. *Proceedings of the National Academy of Sciences of the United States of America*, *121*(32), e2403490121.

Iñesta, C., Oltra-Cucarella, J., Bonete-López, B., Calderón-Rubio, E., & Sitges-Maciá, E. (2021). Regression-based normative data for independent and cognitively active Spanish older adults: Digit span, letters and numbers, trail making test and symbol digit modalities test. *International Journal of Environmental Research & Public Health*, *18*(19), 9958.

Innocenti, F., Tan, F. E. S., Candel, M. J. J. M., & van Breukelen, G. J. P. (2023). Sample size calculation and optimal design for regression based norming of tests and questionnaires. *Psychological Methods*, *28*(1), 89–106.

Instituto Nacional de Estadística y Censos (INDEC) (2024). Estructura de la población y envejecimiento. Buenos Aires, Argentina. https://www.indec.gob.ar/indec/web/Nivel4-Tema-2-21-130.

Ishii, R., Canuet, L., Aoki, Y., Hata, M., Iwase, M., Ikeda, S., Nishida, K., & Ikeda, M. (2017). Healthy and pathological brain aging: From the perspective of oscillations, functional connectivity, and signal complexity. *Neuropsychobiology*, *75*(4), 151–161.

Jawinski, P., Markett, S., Drewelies, J., Düzel, S., Demuth, I., Steinhagen-Thiessen, E., Wagner, G. G., Gerstorf, D., Lindenberger, U., Gaser, C., & Kühn, S. (2022). Linking brain age gap to mental and physical health in the Berlin Aging Study II. *Frontiers in Aging Neuroscience*, *14*, 791222. https://doi.org/10.3389/fnagi.2022.791222

Jokinen, H., Melkas, S., Madureira, S., Verdelho, A., Ferro, Jé M., Fazekas, F., Schmidt, R., Scheltens, P., Barkhof, F., Wardlaw, J. M., Inzitari, D., Pantoni, L., & Erkinjuntti, T. (2016). Cognitive reserve moderates long-term cognitive and functional outcome in cerebral small vessel disease. *Journal of Neurology, Neurosurgery & Psychiatry*, *87*(12), 1296–1302.

Karstens, A. J., Christianson, T. J., Lundt, E. S., Machulda, M. M., Mielke, M. M., Fields, J. A., Kremers, W. K., Graff-Radford, J., Vemuri, P., Jack, C. R.Jr, Knopman, D. S., Petersen, R. C., & Stricker, N. H. (2024). Mayo normative studies: Regression-based normative data for ages 30–91 years with a focus on the Boston naming test, trail making test and category fluency. *Journal of the International Neuropsychological Society: JINS*, *30*(4), 389–401.

Kiselica, A. M., Kaser, A. N., Weitzner, D. S., Mikula, C. M., Boone, A., Woods, S. P., Wolf, T. J., & Webber, T. A. (2024). Development and validity of norms for cognitive dispersion on the uniform data set 3.0 neuropsychological battery. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *39*(6), 732–746. https://doi.org/10.1093/arclin/acae005

Korotkevich, Y., Trewartha, K. M., Penhune, V. B., & Li, K. Z. (2015). Effects of age and cognitive load on response reprogramming. *Experimental Brain Research*, *233*, 937–946.

Leis, A., Allegri, R. F., Román, F., Iturry, M., Crotti, B., Gatto, E., & Rojas, G. J. (2018). Datos normativos de la versión argentina de la batería de eficacia mnésica Signoret (BEM 144) para ser aplicados en la evaluación neurocognitiva. *Neurología Argentina*, *10*(3), 127–136. https://doi.org/10.1016/j.neuarg.2018.04.002

Lezak, M. D. (1983). *Neuropsychological assessment* (2nd ed.). Oxford University Press.

Livingston, G., Huntley, J., Liu, K. Y., Costafreda, S. G., Selbæk, G., Alladi, S., Ames, D., Banerjee, S., Burns, A., Brayne, C., Fox, N. C., Ferri, C. P., Gitlin, L. N., Howard, R., Kales, H. C., Kivimäki, M., Larson, E. B., Nakasujja, N., Rockwood, K., . . . , Mukadam, N. (2024). Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. *Lancet (London, England)*, *404*(10452), 572–628.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 20:33:32, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1355617726101994

Lövdén, M., Fratiglioni, L., Glymour, M. M., Lindenberger, U., & Tucker-Drob, E. M. (2020). Education and cognitive functioning across the life span. *Psychological Science in the Public Interest*, *21*(1), 6–41.

Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health*, *23*, 151–169.

Lundervold, A. J., Wollschlager, D., & Wehling, E. (2014). Age and sex related changes in episodic memory function in middle aged and older adults. *Scandinavian Journal of Psychology*, *55*, 225–232.

Margulis, L. E., Squillace Louhau, M. R., Ferreres, A. R., Margulis, E., Louhau, M. R. Squillace, & Ferreres, A. R. (2019). Baremo del Trail Making Test para Capital Federal y Gran Buenos Aires (L. E. Margulis, M. R. Squillace Louhau, & A. R. Ferreres, Trans.). *Revista Argentina de Ciencias del Comportamiento*, *10*(3), 54–63.

Marino, J., & Alderete, A. M. (2010). Valores normativos de pruebas de fluidez verbal categoriales, fonológicas, gramaticales y combinadas y análisis comparativo de la capacidad de iniciación. *Revista Neuropsicología, Neuropsiquiatría y Neurociencias*, *10*(1), 79–93.

Martino, P., Cervigni, M., Portillo, N., Gallegos, M., Politis, D., Bortoli, M.Á., & Vivas, J. (2022). Three verbal fluency tasks: Normative data and convergent validity in Argentines over 50 years. *Dementia & Neuropsychologia*, *16*(4), 481–488.

Martino, P. L., Cervigni, M. A., Gallegos de San Vicente, M. O., & Politis, D. G. (2021). Estudios normativos argentinos sobre pruebas cognitivas para adultos: Una revisión sistemática (2000-2020). *Revista Argentina de Ciencias del Comportamiento*, *13*(3), 19–33.

Marzuki, A. A., Wong, K. Y., Chan, J. K., Na, S. Y., Thanaraju, A., Phon-Amnuaisuk, P., Vafa, S., Yap, J., Lim, W. G., Yip, W. Z., Arokiaraj, A. S., Shee, D., Lee, L. G. L., Chia, Y. C., Jenkins, M., & Schaefer, A. (2024). Mapping computational cognitive profiles of aging to dissociable brain and sociodemographic factors. *npj Aging*, *10*(1), 50.

Mitrushina, Maura, Boone, Kyle B., Razani, Jill, & D'Elia, Louis F. (2005). *Handbook of normative data for neuropsychological assessment* (2nd ed.). Oxford University Press.

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., Mellits, E. D., & Clark, C. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, *39*(9), 1159–1165. https://doi.org/10.1212/wnl.39.9.1159

Morris, J. C., Weintraub, S., Chui, H. C., Cummings, J., DeCarli, C., Ferris, S., Foster, N. L., Galasko, D., Graff-Radford, N., Peskind, E. R., Beekly, D., Ramos, E. M., & Kukull, W. A. (2006). The Uniform Data Set (UDS): Clinical and cognitive variables and descriptive data from Alzheimer Disease Centers. *Alzheimer Disease and Associated Disorders*, *20*(4), 210–216. https://doi.org/10.1097/01.wad.0000213865.09806.92

Murman, D. L. (2015). The impact of age on cognition. *Seminars in Hearing*, *36*(03), 111–121.

Murty, D. V. P. S., Manikandan, K., Kumar, W. S., Ramesh, R. G., Purokayastha, S., Javali, M., Rao, N. P., & Ray, S. (2020). Gamma oscillations weaken with age in healthy elderly in human EEG. *NeuroImage*, *215*, 116826.

Nichols, E. S., Wild, C. J., Owen, A. M., & Soddu, A. (2021). Cognition across the lifespan: Investigating age, sex, and other sociodemographic influences. *Behavioral Sciences (Basel, Switzerland)*, *11*(4), 51.

Norman, M. A., Evan, J. D., Miller, S. W., & Heaton, R. K. (2000). Demographically corrected norms for the California verbal learning test. *Journal of Clinical & Experimental Neuropsychology*, *22*(1), 80–94.

Norman, M. A., Moore, D. J., Taylor, M., Franklin, D., Jr, Cysique, L., Ake, C., Lazarretto, D., Vaida, F., Heaton, R. K., & HNRC Group (2011). Demographically corrected norms for African Americans and Caucasians on the hopkins verbal learning test-revised, brief visuospatial memory test-revised, stroop color and word test, and wisconsin card sorting test 64-card version. *Journal of Clinical and Experimental Neuropsychology*, *33*(7), 793–804. https://doi.org/10.1080/13803395.2011.559157

Olsson-Collentine, A., Wicherts, J. M., & van Assen, M. A. L. M. (2020). Heterogeneity in direct replications in psychology and its association with effect size. *Psychological Bulletin*, *146*(10), 922–940.

Oosterhuis, H. E., Van der Ark, L. A., & Sijtsma, K. (2016). Sample size requirements for traditional and regression-based norms. *Assessment*, *23*(2), 191–202.

Opdebeeck, C., Martyr, A., & Clare, L. (2016). Cognitive reserve and cognitive function in healthy older people: A meta-analysis. *Aging, Neuropsychology, and Cognition*, *23*(1), 40–60.

Perez-Lao, A., Ying, G., Arias, F., Levy, S. A., & Smith, G. E. (2025). A systematic review of available normative data on neuropsychological tests for Spanish speakers in the U.S., Latin America and the Caribbean, and Spain. *Neuropsychology Review*. Advance online publication. https://doi.org/10.1007/s11065-025-09666-6

Pinaya, W. H. L., Scarpazza, C., Garcia-Dias, R., Vieira, S., Baecker, L., da Costa, P. F., Redolfi, A., Frisoni, G. B., Pievani, M., Calhoun, V. D., Sato, J. R., & Mechelli, A. (2021). Using normative modelling to detect disease progression in mild cognitive impairment and Alzheimer's disease in a cross-sectional multi-cohort study. *Scientific Reports*, *11*(1), 15746. https://doi.org/10.1038/s41598-021-95098-0

Reas, E. T., Laughlin, G. A., Bergstrom, J., Kritz-Silverstein, D., Barrett-Connor, E., & McEvoy, L. K. (2017). Effects of sex and education on cognitive change over a 27-year period in older adults: The Rancho Bernardo study. *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry*, *25*(8), 889–899.

Reitan, R. M. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual & Motor Skills*, *8*(3), 271–276.

Rexroth, D. F., Tennstedt, S. L., Jones, R. N., Guey, L. T., Rebok, G. W., Marsiske, M. M., Xu, Y., & Unverzagt, F. W. (2013). Relationship of demographic and health factors to cognition in older adults in the ACTIVE study. *Journal of Aging & Health*, *25*(8 Suppl), 128S–46S.

Rey, A. (1964). *L'examen clinique en psychologie* (2nd ed.). Presses Universitaires de France.

Richardson, J. T. E. (2007). Measures of short-term memory: A historical review. *Cortex*, *43*(5), 635–650.

Román, F., Iturry, M., Rojas, G., Barceló, E., Buschke, H., & Allegri, R. F. (2016). Validation of the Argentine version of the memory binding test (MBT) for early detection of mild cognitive impairment. *Dementia & Neuropsychologia*, *10*(3), 217–226.

Rossini, P. M., Rossi, S., Babiloni, C., & Polich, J. (2007). Clinical neurophysiology of aging brain: From normal aging to neurodegeneration. *Progress in Neurobiology*, *83*(6), 375–400.

Sachs, B. C., Steenland, K., Zhao, L., Hughes, T. M., Weintraub, S., Dodge, H. H., Barnes, L. L., Craft, S., Parker, M. L., & Goldstein, F. C. (2020). Expanded demographic norms for version 3 of the Alzheimer disease centers neuropsychological test battery in the uniform data set. *Alzheimer Disease and Associated Disorders*, *34*(3), 191–197.

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging*, *30*, 507–514.

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General*, *132*, 566–594.

Schneeweis, N., Skirbekk, V., & Winter-Ebmer, R. (2014). Does education improve cognitive performance four decades after school completion? *Demography*, *51*(2), 619–643, [PubMed: 24578168].

Shirk, S. D., Mitchell, M. B., Shaughnessy, L. W., Sherman, J. C., Locascio, J. J., Weintraub, S., & Atri, A. (2011). A web-based normative calculator for the uniform data set (UDS) neuropsychological test battery. *Alzheimer's Research & Therapy*, *3*(6), 32.

Soer, R., van der Schans, C. P., Geertzen, J. H., Groothoff, J. W., Brouwer, S., Dijkstra, P. U., & Reneman, M. F. (2009). Normative values for a functional capacity evaluation. *Archives of Physical Medicine and Rehabilitation*, *90*(10), 1785–1794. https://doi.org/10.1016/j.apmr.2009.05.008

Sorel, O., & Pennequin, V. (2008). Aging of the planning process: The role of executive functioning. *Brain and Cognition*, *66*, 196–201.

Spreng, R. N., & Turner, G. R. (2019). The shifting architecture of cognition and brain function in older adulthood. *Perspectives on Psychological Science*, *14*, 523–542.

Stacey, J. E., Crook-Rumsey, M., Sumich, A., Howard, C. J., Crawford, T., & Livne, K. 2021). Age differences in resting state EEG and their relation

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 20:33:32, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1355617726101994

to eye movements and cognitive performance'. *Neuropsychologia*, *157*, 107887.

Stier, C., Braun, C., & Focke, N. K. (2023). Adult lifespan trajectories of neuromagnetic signals and interrelations with cortical thickness. *NeuroImage*, *278*, 120275.

Sunderaraman, P., Blumen, H. M., DeMatteo, D., Apa, Z. L., & Cosentino, S. (2013). Task demand influences relationships among sex, clustering strategy, and recall: 16-word versus 9-word list learning tests. *Cognitive and Behavioral Neurology*, *26*, 78–84.

Timmerman, M. E., Voncken, L., & Albers, C. J. (2021). A tutorial on regression-based norming of psychological tests with GAMLSS. *Psychological Methods*, *26*(3), 357–373.

Trammell, J. P., MacRae, P. G., Davis, G., Bergstedt, D., & Anderson, A. E. (2017). The relationship of cognitive performance and the theta-alpha power ratio is age-dependent: An EEG study of short term memory and reasoning during task and resting-state in healthy young and old adults. *Frontiers in Aging Neuroscience*, *9*, 364.

Trondle, M., Popov, T., Pedroni, A., Pfeiffer, C., Baranczuk-Turska, Z., & Langer, N. (2023). Decomposing age effects in EEG alpha power. *Cortex*, *161*, 116–144.

Valdes-Sosa, P. A., Galan-Garcia, L., Bosch-Bayard, J. *et al.* (2021). The Cuban Human Brain Mapping Project, a young and middle age population-based EEG, MRI, and cognition dataset. *Sci Data*, *8*, 45. https://doi.org/10.1038/s41597-021-00829-7

Van Breukelen, G. J. P., & Vlaeyen, J. W. S. (2005). Norming clinical questionnaires with multiple regression: The pain cognition list. *Psychological Assessment*, *17*(3), 336–344.

Van der Elst, W. (2024). *Regression-based normative data for psychological assessment: A hands-on approach using R*. Springer. https://doi.org/10.1007/978-3-031-50951-3.

Van der Elst, W., Molenberghs, G., Van Boxtel, M. P. J., & Jolles, J. (2013). Establishing normative data for repeated cognitive assessment: A comparison of different statistical methods. *Behavior Research Methods*, *45*(4), 1073–1086.

Vaughan, L., & Giovanello, K. (2010). Executive function in daily life: Age-related influences of executive processes on instrumental activities of daily living. *Psychology and Aging*, *25*, 343–355.

Verdi, S., Kia, S. M., Schott, J. M., Marquand, A. F., Cole, J. H., & The Alzheimer's Disease Neuroimaging Initiative. (2022). Mapping individualised patterns of atrophy in Alzheimer's disease using neuroanatomical normative models. *Alzheimer's & Dementia*, *18*, e060306. https://doi.org/10.1002/alz.060306

Voytek, B., Kramer, M. A., Case, J., Lepage, K. Q., Tempesta, Z. R., Knight, R. T., & Gazzaley, A. (2015). Age-related changes in 1/f neural electrophysiological noise. *Journal of Neuroscience*, *35*(38), 13257–13265. https://doi.org/10.1523/JNEUROSCI.2332-14.2015

Weintraub, S. (2009). The Alzheimer's Disease Centers' Uniform Data Set (UDS): The neuropsychologic test battery. *Alzheimer Disease & Associated Disorders*, *23*(2), 91–101.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, *48*(4), 817–838.

Wilson, R. S., Yu, L., Lamar, M., Schneider, J. A., Boyle, P. A., & Bennett, D. A. (2019). Education and cognitive reserve in old age. *Neurology*, *92*(10), e1041–e1050. [PubMed: 30728309].

Yan, J., Li, T., Wang, H., Huang, H., Wan, J., & Nho, K. (2015). Cortical surface biomarkers for predicting cognitive outcomes using group l2,1 norm. *Neurobiology of Aging*, *36*(Suppl 1), S185–S193.

Zachary, R. A., & Gorsuch, R. L. (1985). Continuous norming: Implications for the WAIS-R. *Journal of Clinical Psychology*, *41*(1), 86–94.

Zaninotto, P., Batty, G. D., Allerhand, M., & Deary, I. J. (2018). Cognitive function trajectories and their determinants in older people: 8 years of follow-up in the English longitudinal study of ageing. *Journal of Epidemiology & Community Health*, *72*(8), 685–694.

Zappasodi, F., Marzetti, L., Olejarczyk, E., Tecchio, F., & Pizzella, V. (2015). Age-related changes in electroencephalographic signal complexity. *PLoS One*, *10*(11), e0141995.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 19 Nov 2026 at 20:33:32, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1355617726101994