

Regression-Based Norms for the Hopkins Verbal Learning Test-Revised and the Rey–Osterrieth Complex Figure in a Portuguese Adult Population

Selene G. Vicente¹, Daniela Ramos-Usuga^{2,3}, Fernando Barbosa⁴, Nuno Gaspar¹, Artemisa R. Dorés⁵,
Diego Rivera⁶, Juan Carlos Arango-Lasprilla^{2,7,8,*}

¹Centre for Psychology, Faculty of Psychology and Education Sciences, University of Porto, Porto 4099-002, Portugal

²Biocruces Bizkaia Health Research Institute, Barakaldo 48903, Spain

³Biomedical Research Doctorate Program, University of the Basque Country (UPV/EHU), Leioa 48940, Spain

⁴Laboratory of Neuropsychophysiology, Faculty of Psychology and Education Sciences, University of Porto, Porto 4099-002, Portugal

⁵School of Health, Polytechnic of Porto, Porto 4200-072, Portugal

⁶Departamento de Ciencias de la Salud, Universidad Pública de Navarra, Navarra 31006, España

⁷IKERBASQUE. Basque Foundation for Science, Bilbao 48013, Spain

⁸Department of Cell Biology and Histology, University of the Basque Country (UPV/EHU), Leioa 48940, Spain

*Corresponding author at: Biocruces Bizkaia Health Research Institute, Barakaldo 48903, Spain. Tel.: (34) 946006000 (Ext. 7963).

E-mail address: jcalasprilla@gmail.com (J.C. Arango-Lasprilla).

Received 24 February 2020; revised 16 June 2020; Accepted 16 September 2020

Abstract

Objective: The principal goal of this study was to produce adjusted normative data for European Portuguese native speakers from Portugal on 2 neuropsychological tests widely used to assess learning and memory: the Hopkins Verbal Learning Test-Revised (HVLT-R) and the Rey–Osterrieth Complex Figure Test (ROCF).

Method: The study included 300 individuals aged 18–92 years ($M = 50.4$, $SD = 21.2$), who had educational backgrounds ranging from 3 to 25 years ($M = 10.4$, $SD = 5.2$).

Results: Age, education, and sex were significantly associated with HVLT-R and ROCF performance. These demographic variables accounted for 61% of the variance in HVLT-R total recall, 54% in HVLT-R delayed recall, 18% in HVLT-R recognition, 55% in ROCF copy, and 39% in ROCF immediate recall.

Conclusions: The normative data are presented as regression-based algorithms to adjust direct and derived test scores for age, education, and sex. This study provides a calculator of normative data derived from the results of the regression models.

Keywords: Learning and Memory; Norms; Assessment

Introduction

Learning and memory are closely related constructs. Learning is the acquisition of skill or knowledge, whereas memory is the capacity to store information and use it for adaptive purposes (Anderson, 2000; Baddeley, 2014; Baddeley, Eysenck, & Anderson, 2015; Morgado, 2014). Patients with learning and memory impairments may have problems coding information that prevent them from creating new memories, have difficulties in recovering information, making them unable to access past memories, and even impair prospective memory, altering their ability to remember an action or intention at a future point in time. In all cases, these alterations represent a limitation in various spheres of people's lives, such as emotional and social functioning (Zawadzka & Domańska, 2018), daily life activities (Hollocks, Brookes, Morris, & Markus, 2016), quality of life (Chiaravalloti & DeLuca, 2008; Hollocks et al., 2016), the return to work in long-term disabilities due to neurological and neuropsychiatric disorders (Mitrushina & Tomaszewski, 2019), and even financial management difficulties (Wong et al., 2017). For this reason, it is necessary to carry out an adequate evaluation and diagnosis of these problems, with the aim of being able to establish an intervention plan that allows function to be restored, to stop deterioration, or to compensate for deficits.

Several neuropsychological tests have been developed to evaluate these cognitive processes (Lezak, Howieson, Bigler, & Tranel, 2012; Strauss, Sherman, & Spreen, 2006); however, the two most commonly used tests are the Hopkins Verbal Learning Test-Revised (HVLTR; Brandt & Benedict, 2001) and the Rey–Osterrieth Complex Figure Test (ROCF; Rey, 1941), which have demonstrated good psychometric properties (Cortés, Galindo, & Salvador, 1996; O’Neil-Pirozzi, Goldstein, Strangman, & Glenn, 2012; Tupler, Welsh, Asare-Aboagye, & Dawson, 1995; Woods et al., 2005).

Regarding the influence of sociodemographic variables on performance in these tests, it should be noted that there is a growing body of evidence for the effects of age, sex, education, and ethnicity that must be considered when generating normative data. Studies conducted with the HVLTR and ROCF have found effects of age (Boone, Lesser, Hill-Gutierrez, Berman, & D’elia, 1993; Friedman, Schinka, Mortimer, & Graves, 2002; Hester, Kinsella, Ong, & Turner, 2004; Peña-Casanova et al., 2009; Vanderploeg et al., 2000), sex (Berry, Allen, & Schmitt, 1991; Friedman et al., 2002; Vanderploeg et al., 2000), educational level (Caffarra, Vezzadini, Dieci, Zonato, & Venneri, 2002; Friedman et al., 2002; Hester et al., 2004; Peña-Casanova et al., 2009), and ethnicity (Boone, Victor, Wen, Razani, & Ponton, 2007; Friedman et al., 2002). In addition, given the differences in performance found between people from different languages/countries, normative data for the HVLTR have been developed in countries such as Mexico (Cherner et al., 2007), Australia (Hester et al., 2004; Ryan et al., 2020), the United States (Duff, 2016; Friedman et al., 2002; Ryan et al., 2020), India (Waldrop-Valverde et al., 2015), and Brazil (Miotto et al., 2012). Likewise, the ROCF also has normative data for Italy (Caffarra et al., 2002), Spain (Palomo et al., 2013), Colombia (Rosselli & Ardila, 2003), Denmark (Vogel, Stokholm, & Jorgensen, 2012), Canada (Strauss et al., 2006), New Zealand (Fernando, Chard, Butcher, & McKay, 2003), and the United States (Fastenau, Denburg, & Hufford, 1999).

The majority of these studies have limitations such as different sample sizes with poor control for sociodemographic variables, as well as the use of different methodologies (i.e., the use of mean and standard deviation [SD] within each subgroup or conversion of raw scores to metrics such as *Z* or *T* values). To overcome some of these limitations and to obtain more precise normative data, recently the use of multiple linear regression analyses has been carried out to better understand the influence of sociodemographic factors on test performance. This method provides practical and more reliable normative data by first indicating that sociodemographic factors, or interactions of factors, are significantly associated with performance. Both linear and nonlinear relationships are considered. The regression model thus indicates that sociodemographic groups are required for differentiated normative data tables. This method does not convert raw scores to *Z* scores. Instead, it uses the regression equation to predict an expected test score based on sociodemographic characteristics and the difference between actual and predicted scores is standardized using the residual standard deviation (SDe) value provided by the regression model (Rivera & Arango-Lasprilla, 2017).

Taking into account that to date no normative data have been created for the HVLTR and ROCF in Portugal (there is only a normative study by Cavaco et al., 2015, using the Auditory Verbal Learning Test), the purpose of the present study is to generate normative data for these two neuropsychological tests for native European Portuguese speakers using multiple linear regression analyses.

Material and Methods

Participants

Three-hundred healthy individuals recruited from the Oporto District participated in the present study. Sixty-two percent of the sample were women, the average age was 50.4 ± 21.2 years (range = 18–92), and the average years of education was 10.4 ± 5.2 (range = 3–25). Demographic characteristics of the sample are presented in the Table 1.

Participants who met the following criteria were included: age between 18 and 95 years; were born and currently live in Portugal; have European Portuguese as their native language; had completed at least 1 year of formal education; were literate; scored ≥ 23 on the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975); scored ≤ 4 on the Patient Health Questionnaire–9 (Kroenke, Spitzer, & Williams, 2001); and scored ≥ 90 on the Barthel Index for Activities of Daily Living (Mahoney & Barthel, 1965). Participants with a diagnosis of neurological and/or psychiatric conditions; abuse of an illicit substance; history of systemic disease (e.g., diabetes mellitus); prescription of any type of medication that may alter cognitive functioning; and/or severe visual and/or hearing deficits were excluded.

Instruments and Measures

Self-report questionnaire. A researcher-created questionnaire was used to collect information about participants’ medical history, sociodemographic, and health status.

Table 1. Demographic characteristics of the sample

Age group (years)	N	Age		Education		Sex	
		Mean	SD	Mean	SD	Woman n	Man n
20 ± 2	19	20.7	1.5	12.7	2.1	13	6
25 ± 2	56	24.5	1.3	15.3	2.1	37	19
30 ± 2	16	30.4	1.6	15.3	3.8	11	5
35 ± 2	11	34.8	1.7	13.6	4.7	4	7
40 ± 2	18	40.4	1.4	12.1	4.4	7	11
45 ± 2	13	44.7	1.5	11.4	5.5	6	7
50 ± 2	28	50.2	1.2	9.6	4.8	15	13
55 ± 2	11	54.9	1.8	11.2	4.8	9	2
60 ± 2	21	59.8	1.5	9.4	4.3	15	6
65 ± 2	28	65.4	1.3	7.2	4.0	20	8
70 ± 2	22	69.4	1.5	7.8	4.4	16	6
75 ± 2	26	75.0	1.6	5.7	3.6	14	12
80 ± 2	12	79.3	1.6	4.3	1.5	8	4
>82	19	87.0	3.0	4.6	2.2	13	6
Total	300	50.4	21.2	10.4	5.2	188	112

The Hopkins Verbal Learning Test-Revised. The test consists of three parts: total recall, delayed recall, and recognition. HVLTR total recall consists of three free recall learning trials with 12 words from three categories (professions, food, and sports). Twenty minutes later, the participant is asked to name the words from the free recall learning list (HVLTR delayed recall). Finally, the yes/no recognition trial is completed (HVLTR recognition). This last task has a total of 24 words: the 12 words from the recall list, another 6 words semantically related to the recall items, and 6 unrelated words (Brandt & Benedict, 2001).

Rey–Osterrieth Complex Figure Test. The test is used to evaluate perceptual organization and visual memory, assessing the ability to organize and plan strategies for problem solving, as well as visuoconstructive ability. In the first part, the figure is presented, and the participant is prompted to copy it (ROCF Copy). Three minutes after the copy is finished, the participant is asked to perform the figure again without the stimulus (ROCF immediate recall). There is no time limit for any of the tasks. The ROCF is composed of 18 elements, and two points are given when the elements are correctly reproduced, one point when the reproduction is distorted, incomplete but placed properly, or completed but placed poorly, and a half point is credited when the elements are distorted or incomplete and placed poorly. A zero score is given when the elements are absent or not recognizable. The maximum score for each task is 36 (Osterrieth, 1944).

Procedure

Data collection took place over a period of 6 months in the northern region of Portugal (Oporto district). The HVLTR and the ROCF were administered in Portuguese by two Portuguese neuropsychologists as a part of a comprehensive neuropsychological battery in a single day, lasting about 70 min. Before test administration, each participant completed and signed an informed consent, according to the Helsinki Declaration. Participation in the study was completely voluntary without any monetary compensation. This study was approved by the Ethics Committee of the University of Oporto.

Statistical Analyses

Exploratory data analysis. Pearson correlations between the HVLTR (total recall, delayed recall, and recognition) and the ROCF (copy and immediate recall) scores, and the sociodemographic variables (age, education, and sex) were computed.

Demographic effects and generation of normative data. Normative data based on multiple linear regression analyses were used to determine the effects of demographic variables on each score separately (HVLTR total recall, HVLTR delayed recall, HVLTR recognition, ROCF copy, and ROCF immediate recall). Age, education (as continuous variables), and sex were used as predictors of performance. Moreover, for age and education, polynomial second order trends (age², education²) were also

Table 2. Standard deviation (residual) for final multiple linear regression models

Test	Predicted value	SD
HVLT-R Total Recall	All values	3.584
HVLT-R Delayed	All values	1.973
HVLT-R Recognition	≤10.546	1.507
	10.547–11.106	1.216
	11.107–11.435	1.034
	≥11.436	0.608
ROCF Copy	≤26.065	6.252
	26.066–31.589	4.242
	31.590–33.724	4.195
	≥33.725	1.923
ROCF Immediate Recall	All values	5.667

Note: HVLT-R = Hopkins Verbal Learning Test-Revised; ROCF = Rey–Osterrieth Complex Figure.

included in the model in order to assess curvilinear effects, centralizing them first to reduce multicollinearity (age in years— $\bar{x}$ age in the sample; education in years— $\bar{x}$ education in the sample). Exploring more than one type of function (linear vs. quadratic) for age and education is a primary advantage of regression-based normative data (Van Breukelen & Vlaeyen, 2005). This method allows the determination of the best model to explain each neuropsychological score. Sex was coded as woman = 0 and man = 1. Finally, all two-way interactions between age, age², education, education², and sex were also included. The model assumes that the residuals ε_i are normally distributed with mean 0 and variance σ_ε^2 , that is, $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$.

Variables that were not statistically significant (Bonferroni-corrected α -level of .01) were removed stepwise, and a truncated model was run again, until the final model was obtained. During this process, predictors that were also included in a higher order term in the multiple regression model (quadratic variables and interactions) were not removed (Aiken, West, & Reno, 1991). Then, four assumptions were evaluated for each model: multicollinearity (Variance Inflation Factor [VIF] ≤ 10), homoscedasticity (participants were grouped into quartiles of the predicted scores and the Levene's test was applied on the residuals), normality of the standardized residuals (Kolmogorov–Smirnov test), and the existence of influential values assessed [calculation of the maximum Cook's distance and subsequently related to an $F(p, n - p)$ distribution] (Kutner, Nachtsheim, Neter, & Li, 2005).

Finally, a four-step procedure was carried out for HVLT-R and ROCF scores separately to generate normative data adjusted by sociodemographic variables (Rivera et al., 2019; Van Breukelen & Vlaeyen, 2005; Van der Elst, van Boxtel, Van Breukelen, & Jolles, 2006a, 2006b): The expected test score ($\hat{Y}_i$) is computed using fixed effect parameter from the final regression model: $\hat{Y}_i = B_0 + B_1X_{1i} + B_2X_{2i} + \dots + B_KX_{Ki}$; the residual value $e_i = Y_i - \hat{Y}_i$ is calculated; the residuals are standardized using the residual SD (SD_e) value provided by the regression model (Table 2): $z_i = e_i/SD_e$; and the exact percentile corresponding to the z -score is obtained using the standard normal cumulative distribution function (if normality assumption is met) or the empirical cumulative distribution function of the standardized residuals (if the normality assumption is not met).

Adjusted R^2 values are provided for all final models. SPSS Version 23 (IBM Corp., Armonk, NY) and R 3.4.2 for Windows (R Development Core Team, 2017) were used to perform the analyses.

Results

Exploratory Data Analysis

The intercorrelations between age and HVLT-R (total recall, delayed recall, and recognition) and ROCF (copy and immediate recall) scores were significantly negative (all $r \geq -.375$, all $p < .001$) and significantly positive with education (all $r \geq .357$, $p < .001$). Otherwise, HVLT-R recognition and ROCF copy and immediate recall scores correlated significantly negative with sex (all $r \geq -.0121$, $p < .05$; Table 3).

Model Assumptions

The assumptions of multiple regression analysis show there was no multicollinearity all final models (VIF values ≤ 2.599) or influential cases [maximum Cook's distance = 0.156; relating this value to an $F_{(5,295)}$ distribution yields a percentile value of 2, which is below the threshold (Percentile = 50) that is indicative for the presence of influential cases], neither homoscedasticity

Table 3. Correlations between all HVLTR and ROCF scores and demographic variables

	Age	Education	Sex
HVLT-R Total Recall	-.665**	.623**	.042
HVLT-R Delayed	-.576**	.585**	.006
HVLT-R Recognition	-.722**	.690**	-.121*
ROCF Copy	-.670**	.662**	-.126*
ROCF Immediate Recall	-.375**	.357**	-.138*

Note: HVLTR = Hopkins Verbal Learning Test-Revised; ROCF = Rey–Osterrieth Complex Figure.

** $p < .001$

* $p < .05$

Table 4. Final multiple linear regression models for HVLTR scores

Score		<i>B</i>	<i>SE</i>	β	<i>t</i>	Sig.	Adj <i>R</i> ²
Total Recall	(Intercept)	23.136	.368		62.844	<.001	.612
	Age	-.121	.014	-.443	-8.521	<.001	
	Education	.429	.058	.386	7.425	<.001	
	Education ²	-.029	.009	-.131	-3.440	.001	
	Sex	-1.491	.441	-.125	-3.380	.001	
Delayed	(Intercept)	7.884	.203		38.890	<.001	.545
	Age	-.053	.008	-.382	-6.791	<.001	
	Education	.226	.032	.399	7.090	<.001	
	Education ²	-.015	.005	-.132	-3.207	.001	
	Sex	-.772	.243	-.127	-3.176	.002	
Recognition	(Intercept)	11.366	.116		97.587	<.001	.184
	Age	-.013	.005	-.220	-2.918	.004	
	Education	.050	.018	.206	2.734	.007	
	Education ²	-.007	.003	-.146	-2.651	.008	
	Sex	-.400	.140	-.153	-2.867	.004	

Note: HVLTR = Hopkins Verbal Learning Test-Revised; SE = standard error.

in all models except in HVLTR recognition and ROCF immediate recall scores. Standardized residuals of the models were normally distributed except for HVLTR recognition and ROCF immediate recall.

The Effects of Demographic Variables

Hopkins Verbal Learning Test-Revised. The final multiple linear regression models for HVLTR total recall, delayed, and recognition scores were significant (Table 4). HVLTR total recall, delayed recall, and recognition scores were negatively influenced by age, quadratic education, and sex, so young women best performed the task. The performance showed a curvilinear pattern according to education, such that the scores increased until 13 years of education in all ages and remained stable after 14 years of education (Fig. 1). The amount of variance explained by these predictors was 61% for HVLTR total recall, 54% for HVLTR delayed recall, and 18% for HVLTR recognition scores.

Rey–Osterrieth Complex Figure Test. The final multiple linear regression models for ROCF copy and ROCF immediate recall scores were significant (Table 5). ROCF copy score decreased curvilinearly as a function of quadratic age and quadratic education. In this way, the scores remained the same until the age of 45. At that point, scores begin to decrease showing a curvilinear pattern. Education also affected the scores differently by sex, such that women's performance increased curvilinearly until 14 years of education and then remained stable. On the contrary, men's scores increased until 10 years of education, remained stable between 10 and 14 years, and started to decrease with more than 14 years of education (Fig. 2). ROCF immediate recall score was negatively affected by age and increased linearly as a function of education, showing that young people with higher educational levels had better performance. The amount of variance explained by these predictors for ROCF copy and ROCF immediate recall scores was 55% and 39%, respectively.

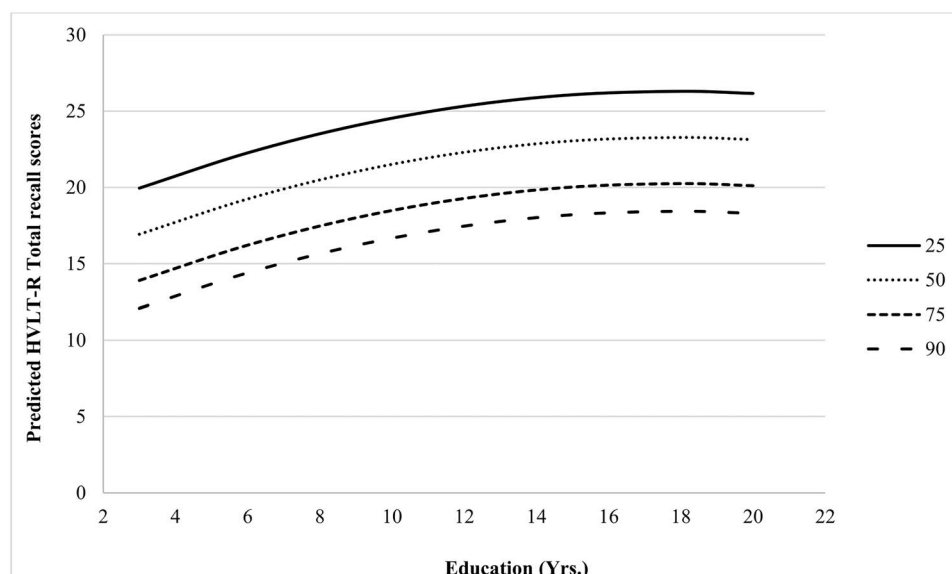


Fig. 1. Predicted mean scores as a function of age and education for HVLTR Total recall scores from a Portuguese sample.

Table 5. Final multiple linear regression models for ROCF scores

Score		<i>B</i>	<i>SE</i>	β	<i>t</i>	Sig.	Adj <i>R</i> ²
Copy	(Intercept)	31.944	.540		59.157	<.001	.551
	Age	-.116	.018	-.370	-6.615	<.001	
	Age ²	-.002	.001	-.132	-3.348	.001	
	Education	.588	.080	.461	7.379	<.001	
	Education ²	-.047	.011	-.181	-4.358	<.001	
	Sex	.127	.549	.009	.231	.818	
	Edu × Sex	-.389	.111	-.163	-3.494	.001	
Immediate Recall	(Intercept)	15.968	.328		48.635	<.001	.395
	Age	-.114	.022	-.330	-5.290	<.001	
	Education	.500	.087	.357	5.725	<.001	

Note: ROCF = Rey–Osterrieth Complex Figure.

Normative Data Calculator

The procedure described previously is long and requires complex operations, increasing the chance for mistakes. For this reason, a calculator in Microsoft Excel was developed in order to obtain the Z scores and the exact percentiles by adding the raw scores and demographic variables (age, sex, and education). This tool is freely available for all users and can be downloaded at <https://neuropsychologylearning.com/datos-normativos-archivos-descargables/>.

Discussion

This study aimed to generate adjusted normative data for the HVLTR and the ROCF in a Portuguese healthy adult sample. A regression-based normative method was used including age, education, and sex in order to determine the influence of sociodemographic variables on test performance.

The final multiple linear regression models explained 61% of the variance in HVLTR total recall score, 54% of the variance in HVLTR delayed recall score, and 18% of the variance in HVLTR recognition score. In the ROCF, the models explained 55% of the variance in copy score and 39% of the variance in immediate recall score. An effect of age and education was found on all scores, whereas sex influenced all scores except for ROCF immediate recall. In all cases, women performed better than men.

These results are consistent with previous studies that have generated normative data for the HVLTR and ROCF. First, there is robust evidence that as age increases, scores get worse (Chervinsky, Mitrushina, & Satz, 1992; Chiulli, Haaland, LaRue, & Garry, 1995; Friedman et al., 2002; Rosselli & Ardila, 1991; Strauss et al., 2006; Vanderploeg et al., 2000). However, it

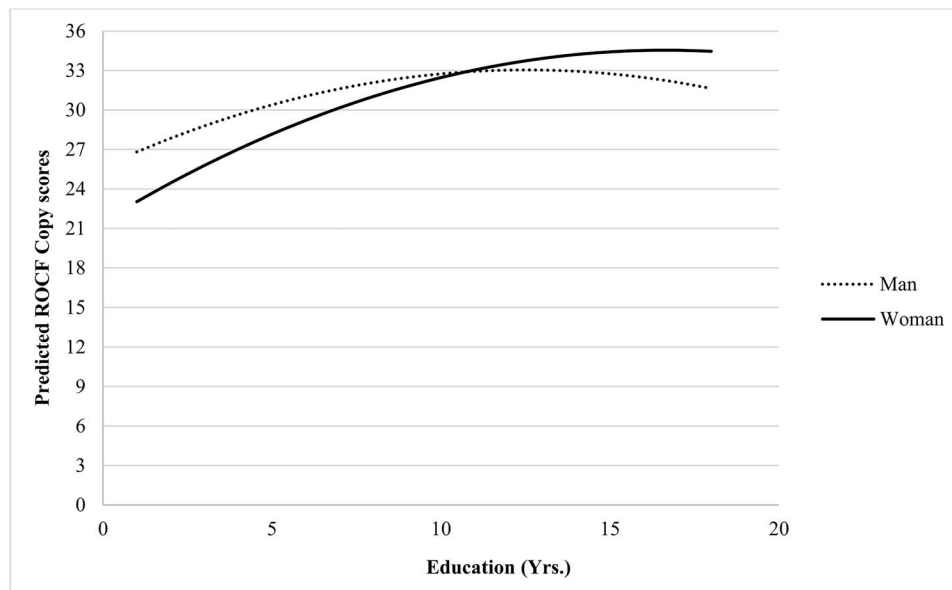


Fig. 2. Predicted expected ROCF Copy scores as a function of education and sex. Predicted expected test scores were presented to show the fitted regression model $\hat{y}_i = 31.944 + [-.116 \cdot (\text{Age} - 50.4)] + [(-.002 \cdot (\text{Age} - 50.4)^2)] + [.588 \cdot (\text{Education}_i - 10.4)] + [(-.047 \cdot (\text{Education}_i - 10.4)^2)] + (.127 \cdot \text{Sex}_i) + [(-.389 \cdot (\text{Education}_i - 10.4) \cdot \text{Sex}_i)]$.

is necessary to emphasize that in some cases there is a curvilinear effect of age, such that scores do not decline linearly but remain stable until a certain age before declining. This was only observed in ROCF copy performance, where the decline began at 45 years old. Likewise, education played an essential role in test performance, such that people with a higher educational level (with more than 12 years of education) had better performance than those with a low educational level, findings that are consistent with the literature (Arango-Lasprilla et al., 2015; Ardila, Rosselli, & Rosas, 1989; Berry et al., 1991; Caffarra et al., 2002; Friedman et al., 2002). In the same way, the curvilinear effect of education has also been observed, such that the scores increased with the number of years of education to finally stabilize at a certain point, resulting in a ceiling effect (Fig. 1). In the present study, all scores except ROCF immediate recall had a curvilinear effect. Finally, the literature has reported that women have better scores than men on the HVLTR (Friedman et al., 2002; Vanderploeg et al., 2000), whereas men outperform women on the ROCF (Berry et al., 1991; Boone et al., 1993; Peña-Casanova et al., 2009). This seems to be because women are better at verbal tasks, and therefore seem to have an advantage in verbal memory tests. On the other hand, men have better visuospatial skills than women that allow them to better perform the memory tests requiring these skills, as is the case with the ROCF (Lewin, Wolgers, & Herlitz, 2001). Although the data from the present study confirm these findings in the case of the HVLTR, it was not confirmed on the ROCF because women performed better than men. It should be highlighted that although significant sex differences have been found in ROCF performance in previous studies, the effect was smaller (Rivera et al., 2015).

The results of the present study have several implications for clinical practice and research. In Portugal, neuropsychology is a young field. In several countries where the discipline is in its early development, neuropsychologists commonly use normative data from other countries due to the lack of adequate norms for their population (Arango-Lasprilla et al., 2017; Branco Lopes et al., 2019; Olabarrieta-Landa et al., 2016; Truter, Mazabow, Morlett Paredes, Rivera, & Arango-Lasprilla, 2018). This practice can be problematic, as research has shown that people from different countries perform differently on neuropsychological tests, even though people may share similar characteristics (Buré-Reyes et al., 2013; Duggan, Awakon, Loaiza, & Garcia-Barrera, 2019). The use of normative data from a different country can lead to a misinterpretation of test results. The following is an example to illustrate the different percentiles that could be obtained using normative data from various countries on the ROCF. A 30-year-old Portuguese woman with 11 years of education gets a score of 19 on the ROCF Immediate Recall. Using the normative data from Fastenau et al. (1999) from North Americans, the raw score would correspond to the 60th–70th percentile. However, if normative data developed for Spain (Palomo et al., 2013) were used, then the score would be placed in the 29th–40th percentile. In contrast, the normative data from the present study in a Portuguese sample place the score in the 52nd percentile. It is advantageous for Portuguese neuropsychologists, and all neuropsychologists, to use normative data that are most appropriate to the characteristics of their population in order to improve the accuracy of the assessment of the cognitive functioning, diagnostic, and treatment planning in neurorehabilitation.

The results of this study must be considered in light of the following limitations. First, the HVLT-R has six forms, each one with different words, which reduces the effect of learning when performing revaluations. However, in the present normative study, data have been generated for only one form. Although the authors of the test report that the forms are very similar in their psychometric properties (Brandt, 1991), it is not appropriate to use these normative data with the other forms. Second, the sample was collected only in the northern region of Portugal (Oporto district) and mainly in urban areas, so it is not representative of rural areas. Future research should consider potential differences in performance based on urban versus rural residence that might be related to factors other than age and years of education, for instance, the quality of the education. Third, the participants included in the study were healthy and literate people; thus, the norms cannot be applied to those who are illiterate.

Despite the limitations, this is the first time that normative data for HVLT-R and ROCF have been developed in Portugal. These are two of the most widely used tests in neuropsychology worldwide, so it is essential to have appropriate normative data for each language and country in order to carry out an appropriate neuropsychological assessment. Moreover, they are very useful tools in clinical neuropsychology for the assessment of different aspects of learning and memory due to their sensitivity to cognitive dysfunction. In addition, this study provides a calculator based on a sophisticated method to generate normative data controlling for sociodemographic variables that influence performance on these tests.

Acknowledgements

Ramos-Usuga, D. was supported by a predoctoral fellowship from the Basque Government (PRE_2019_1_0164).

Funding

Supported by the Neuropsychology Counselling Service, Laboratory of Neuropsychophysiology, and Centre for Psychology of the Faculty of Psychology and Educational Sciences of the University of Porto, Porto, Portugal. This financial support was used to pay the two psychologists for data collection and to buy the materials.

Conflict of Interest

None declared.

References

- Aiken, L. S., West, S. G., & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*. New York: Sage.
- Anderson, J. R. (2000). *Learning and Memory: An integrated approach* (2nd ed.). New York: Wiley.
- Arango-Lasprilla, J. C., Rivera, D., Garza, M. T., Saracho, C. P., Rodríguez, W., Rodríguez-Agudelo, Y. et al. (2015). Hopkins verbal learning test–revised: Normative data for the Latin American Spanish speaking adult population. *NeuroRehabilitation*, 37(4), 699–718. doi: [10.3233/NRE-151286](https://doi.org/10.3233/NRE-151286).
- 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.
- Ardila, A., Rosselli, M., & Rosas, P. (1989). Neuropsychological assessment in illiterates: Visuospatial and memory abilities. *Brain and Cognition*, 11(2), 147–166. doi: [10.1016/0278-2626\(89\)90015-8](https://doi.org/10.1016/0278-2626(89)90015-8).
- Baddeley, A. (2014). *Essentials of Human Memory*. London: Original work published 1999 Psychology Press.
- Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2015). *Memory* (2nd ed.). New York: Psychology Press.
- Berry, D. T., Allen, R. S., & Schmitt, F. A. (1991). Rey-Osterrieth complex figure: Psychometric characteristics in a geriatric sample. *The Clinical Neuropsychologist*, 5(2), 143–153. doi: [10.1080/13854049108403298](https://doi.org/10.1080/13854049108403298).
- Boone, K. B., Lesser, I. M., Hill-Gutierrez, E., Berman, N. G., & D'elia, L. F. (1993). Rey-Osterrieth Complex Figure performance in healthy, older adults: Relationship to age, education, sex, and IQ. *The Clinical Neuropsychologist*, 7(1), 22–28. doi: [10.1080/13854049308401884](https://doi.org/10.1080/13854049308401884).
- Boone, K. B., Victor, T. L., Wen, J., Razani, J., & Pontón, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of Clinical Neuropsychology*, 22(3), 355–365. doi: [10.1016/j.acn.2007.01.010](https://doi.org/10.1016/j.acn.2007.01.010).
- Branco Lopes, A., Leal, G., Malvy, L., Wauquiez, G., Ponchel, A., Rivera, D. et al. (2019). Neuropsychology in France. *Applied Neuropsychology: Adult*, 1–12.
- Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *The Clinical Neuropsychologist*, 5(2), 125–142. doi: [10.1080/13854049108403297](https://doi.org/10.1080/13854049108403297).
- Brandt, J., & Benedict, R. H. B. (2001). *Hopkins Verbal Learning Test–Revised*. Odessa, FL: PAR.
- Buré-Reyes, A., Hidalgo-Ruzzante, N., Vilar-López, R., Gontier, J., Sánchez, L., Pérez-García, M. et al. (2013). Neuropsychological test performance of Spanish speakers: Is performance different across different Spanish-speaking subgroups? *Journal of Clinical and Experimental Neuropsychology*. doi: [10.1080/13803395.2013.778232](https://doi.org/10.1080/13803395.2013.778232).
- Caffarra, P., Vezzadini, G., Dieci, F., Zonato, F., & Venneri, A. (2002). Rey-Osterrieth complex figure: Normative values in an Italian population sample. *Neurological Sciences*, 22(6), 443–447. doi: [10.1007/s100720200003](https://doi.org/10.1007/s100720200003).
- Cavaco, S., Gonçalves, A., Pinto, C., Almeida, E., Gomes, F., Moreira, I. et al. (2015). Auditory Learning Verbal Test in a large nonclinical Portuguese population. *Applied Neuropsychology: Adult*, 22, 321–331. doi: [10.1080/23279095.2014.927767](https://doi.org/10.1080/23279095.2014.927767).

- Cherner, M. M., Suarez, P. P., Lazzaretto, D. D., Fortuny, L., Mindt, M., Dawes, S. S. et al. (2007). Demographically corrected norms for the Brief Visuospatial Memory Test-revised and Hopkins Verbal Learning Test-revised in monolingual Spanish speakers from the U.S.–Mexico border region. *Archives of Clinical Neuropsychology*, 22(3), 343–353. doi: [10.1016/j.acn.2007.01.009](https://doi.org/10.1016/j.acn.2007.01.009).
- Chervinsky, A. B., Mitrushina, M., & Satz, P. (1992). Comparison of four methods of scoring the Rey-Osterrieth Complex Figure Drawing Test on four age groups of normal elderly. *Brain Dysfunction*, 5, 267–287.
- Chiaravalloti, N. D., & DeLuca, J. (2008). Cognitive impairment in multiple sclerosis. *The Lancet Neurology*, 7(12), 1139–1151. doi: [10.1016/S1474-4422\(08\)70259-X](https://doi.org/10.1016/S1474-4422(08)70259-X).
- Chiulli, S. J., Haaland, K. Y., LaRue, A., & Garry, P. (1995). Impact of age in drawing the Rey-Osterrieth figure. *The Clinical Neurophysiologist*, 9(3), 219–224. doi: [10.1080/13854049508400483](https://doi.org/10.1080/13854049508400483).
- Cortés, J. F., Galindo, G., Villa, M., & Salvador, J. (1996). La figura compleja de Rey: Propiedades psicométricas. *Salud Mental*, 19(3), 42–48.
- Duff, K. (2016). Demographically corrected normative data for the Hopkins verbal learning test-revised and brief visuospatial memory test-revised in an elderly sample. *Applied Neuropsychology: Adult*, 23(3), 179–185. doi: [10.1080/23279095.2015.1030019](https://doi.org/10.1080/23279095.2015.1030019).
- Duggan, E. C., Awakon, L. M., Loaiza, C. C., & Garcia-Barrera, M. A. (2019). Contributing towards a cultural neuropsychology assessment decision-making framework: Comparison of WAIS-IV norms from Colombia, Chile, Mexico, Spain, United States, and Canada. *Archives of Clinical Neuropsychology*. doi: [10.1093/arclin/acy074](https://doi.org/10.1093/arclin/acy074).
- Fastenau, P. S., Denburg, N. L., & Hufford, B. J. (1999). Adult norms for the Rey-Osterrieth Complex Figure Test and for supplemental recognition and matching trials from the Extended Complex Figure Test. *The Clinical Neuropsychologist*, 13(1), 30–47.
- Fernando, K., Chard, L., Butcher, M., & McKay, C. (2003). Standardisation of the Rey Complex Figure Test in New Zealand children and adolescents. *New Zealand Journal of Psychology*, 32(1), 33–38.
- 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(3), 189–198. doi: [10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6).
- Friedman, M. A., Schinka, J. A., Mortimer, J. A., & Graves, A. B. (2002). Hopkins Verbal Learning Test—Revised: Norms for elderly African Americans. *The Clinical Neuropsychologist*, 16(3), 356–373. doi: [10.1076/clin.16.3.356.13857](https://doi.org/10.1076/clin.16.3.356.13857).
- Hester, R. L., Kinsella, G. J., Ong, B., & Turner, M. (2004). Hopkins verbal learning test: Normative data for older Australian adults. *Australian Psychologist*, 39(3), 251–255. doi: [10.1080/00050060412331295063](https://doi.org/10.1080/00050060412331295063).
- Hollocks, M. J., Brookes, R., Morris, R. G., & Markus, H. S. (2016). Associations between the Brief Memory and Executive Test (BMET), activities of daily living, and quality of life in patients with cerebral small vessel disease. *Journal of the International Neuropsychological Society*, 22(5), 561–569. doi: [10.1017/S1355617716000187](https://doi.org/10.1017/S1355617716000187).
- Corp, I. B. M. (2015). *IBM SPSS Statistics for Windows, Version 23.0*. Armonk, NY: IBM Corp.
- 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(9), 606–613. doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x).
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models* (5th ed.). New York: McGraw Hill.
- Lewin, C., Wolgers, G., & Herlitz, A. (2001). Sex differences favoring women in verbal but not in visuospatial episodic memory. *Neuropsychology*, 15(2), 165–173. doi: [10.1037//0894-4105.15.2.165](https://doi.org/10.1037//0894-4105.15.2.165).
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). London: Oxford University Press.
- Mahoney, F. I., & Barthel, D. (1965). Functional evaluation: The Barthel index. *Maryland State Medical Journal*, 14, 56–61.
- Miotto, E. C., Campanholo, K. R., Rodrigues, M. M., Serrao, V. T., Lucia, M., & Scaff, M. (2012). Hopkins verbal learning test-revised and brief visuospatial memory test-revised: Preliminary normative data for the Brazilian population. *Arquivos de Neuro-psiquiatria*, 70(12), 962–965. doi: [10.1590/S0004-282X2012001200014](https://doi.org/10.1590/S0004-282X2012001200014).
- Mitrushina, M., & Tomaszewski, R. (2019). Factors associated with return to work in patients with long-term disabilities due to neurological and neuropsychiatric disorders. *Neuropsychological Rehabilitation*, 29(9), 1313–1331. doi: [10.1147/mf.1416200368](https://doi.org/10.1147/mf.1416200368).
- Morgado, I. (2014). *Aprender, recordar y olvidar*. Madrid, España: Editorial Ariel.
- O’Neil-Pirozzi, T. M., Goldstein, R., Strangman, G. E., & Glenn, M. B. (2012). Test–re-test reliability of the Hopkins Verbal Learning Test-Revised in individuals with traumatic brain injury. *Brain Injury*, 26(12), 1425–1430.
- Olabarrieta-Landa, L., Caracul, A., Pérez-García, M., Panyavin, I., Morlett-Paredes, A., & Arango-Lasprilla, J. C. (2016). The profession of neuropsychology in Spain: Results of a national survey. *The Clinical Neuropsychologist*, 30(8), 1335–1355.
- Osterrieth, P. A. (1944). Le test de copie d’une figure complexe: Contribution à l’étude de la perception et la mémoire. *Archives de Psychologie*, 30, 286–356.
- Palomo, R., Casals-Coll, M., Sánchez-Benavides, G., Quintana, M., Manero, R. M., Rognoni, T. et al. (2013). Spanish normative studies in young adults (NEURONORMA young adults project): Norms for the Rey–Osterrieth complex figure (copy and memory) and Free and Cued Selective Reminding Test. *Neurología (English Edition)*, 28(4), 226–235. doi: [10.1016/j.nrl.2012.03.008](https://doi.org/10.1016/j.nrl.2012.03.008).
- Peña-Casanova, J., Gramunt-Fombuena, N., Quiñones-Úbeda, S., Sánchez-Benavides, G., Aguilar, M., Badenes, D. et al. (2009). Spanish multicenter normative studies (NEURONORMA Project): Norms for the Rey–Osterrieth complex figure (copy and memory), and free and cued selective reminding test. *Archives of Clinical Neuropsychology*, 24(4), 371–393. doi: [10.1093/arclin/acp041](https://doi.org/10.1093/arclin/acp041).
- R Development Core Team (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org/>
- Rey, A. (1941). L’examen psychologique dans les cas d’encéphalopathie traumatique. *Archives de Psychologie*, 28, 286–340.
- Rivera, D., & Arango-Lasprilla, J. C. (2017). Methodology for the development of normative data for Spanish-speaking pediatric populations. *NeuroRehabilitation*, 41(3), 581–592. doi: [10.3233/NRE-172275](https://doi.org/10.3233/NRE-172275).
- Rivera, D., Olabarrieta-Landa, L., Van der Elst, W., Gonzalez, I., Rodríguez-Agudelo, Y., Aguayo Arelis, A. et al. (2019). Normative data for verbal fluency in healthy Latin American adults: Letter M, and fruits and occupations categories. *Neuropsychology*, 33(3), 287. doi: [10.1037/neu0000518](https://doi.org/10.1037/neu0000518).
- Rivera, D., Perrin, P. B., Morlett-Paredes, A., Galarza-del-Angel, J., Martinez, C., Garza, M. T. et al. (2015). Rey–Osterrieth Complex Figure–copy and immediate recall: Normative data for the Latin American Spanish speaking adult population. *NeuroRehabilitation*, 37(4), 677–698. doi: [10.3233/NRE-151285](https://doi.org/10.3233/NRE-151285).
- Rosselli, M., & Ardila, A. (1991). Effects of age, education, and gender on the Rey–Osterrieth Complex Figure. *The Clinical Neuropsychologist*, 5(4), 371–376. doi: [10.1080/13854049108404104](https://doi.org/10.1080/13854049108404104).

- Rosselli, M., & Ardila, A. (2003). The impact of culture and education on nonverbal neuropsychological measurements: A critical review. *Brain and Cognition*, 52(3), 326–333. doi: [10.1016/s0278-2626\(03\)00170-2](https://doi.org/10.1016/s0278-2626(03)00170-2).
- Ryan, J., Woods, R. L., Murray, A. M., Shah, R. C., Britt, C. J., Reid, C. M. et al. (2020). Normative performance of older individuals on the Hopkins Verbal Learning Test-Revised (HVLTR) according to ethno-racial group, gender, age and education level. *The Clinical Neuropsychologist*, 1–17. doi: [10.1080/13854046.2020.1730444](https://doi.org/10.1080/13854046.2020.1730444).
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests. Administration, norms, and commentary*. New York: Oxford University.
- Truter, S., Mazabow, M., Morlett Paredes, A., Rivera, D., & Arango-Lasprilla, J. C. (2018). Neuropsychology in South Africa. *Applied Neuropsychology: Adult*, 25(4), 344–355.
- Tupler, L. A., Welsh, K. A., Asare-Aboagye, Y., & Dawson, D. V. (1995). Reliability of the Rey-Osterrieth Complex Figure in use with memory-impaired patients. *Journal of Clinical and Experimental Neuropsychology*, 17(4), 566–579.
- Van Breukelen, G. J., & Vlaeyen, J. W. (2005). Norming clinical questionnaires with multiple regression: The Pain Cognition List. *Psychological Assessment*, 17(3), 336. doi: [10.1037/1040-3590.17.3.336](https://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(01), 80–89. doi: [10.1017/S1355617706060115](https://doi.org/10.1017/S1355617706060115).
- Van der Elst, W., van Boxtel, M. P., van Breukelen, G. J., & Jolles, J. (2006b). The Letter Digit Substitution Test: Normative data for 1,858 healthy participants aged 24–81 from the Maastricht Aging Study (MAAS): Influence of age, education, and sex. *Journal of Clinical and Experimental Neuropsychology*, 28(6), 998–1009. doi: [10.1080/13803390591004428](https://doi.org/10.1080/13803390591004428).
- Vanderploeg, R. D., Schinka, J. A., Jones, T., Small, B. J., Borenstein Graves, A., & Mortimer, J. A. (2000). Elderly norms for the Hopkins verbal learning test-revised. *The Clinical Neuropsychologist*, 14(3), 318–324. doi: [10.1076/1385-4046\(200008\)14:3;1-P;FT318](https://doi.org/10.1076/1385-4046(200008)14:3;1-P;FT318).
- Vogel, A., Stokholm, J., & Jorgensen, K. (2012). Performances on Rey Auditory Verbal Learning Test and Rey Complex Figure Test in a healthy, elderly Danish sample – Reference data and validity issues. *Scandinavian Journal of Psychology*, 53(1), 26–31. doi: [10.1111/j.1467-9450.2011.00909.x](https://doi.org/10.1111/j.1467-9450.2011.00909.x).
- Waldrop-Valverde, D., Ownby, R. L., Jones, D. L., Sharma, S., Nehra, R., Kumar, A. M. et al. (2015). Neuropsychological test performance among healthy persons in northern India: Development of normative data. *Journal of Neurovirology*, 21(4), 433–438. doi: [10.1007/s13365-015-0332-4](https://doi.org/10.1007/s13365-015-0332-4).
- Wong, S., Irish, M., O'Callaghan, C., Kumfor, F., Savage, G., Hodges, J. R. et al. (2017). Should I trust you? Learning and memory of social interactions in dementia. *Neuropsychologia*, 104, 157–167. doi: [10.1016/j.neuropsychologia.2017.08.016](https://doi.org/10.1016/j.neuropsychologia.2017.08.016).
- Woods, S. P., Scott, J. C., Conover, E., Marcotte, T. D., Heaton, R. K., Grant, I. et al. (2005). Test-retest reliability of component process variables within the Hopkins Verbal Learning Test-Revised. *Assessment*, 12(1), 96–100.
- Zawadzka, E., & Domańska, Ł. (2018). Emotional and social characteristics of stroke patients with low verbal memory. *Aging Clinical and Experimental Research*, 30(10), 1203–1210. doi: [10.1007/s40520-018-0894-0](https://doi.org/10.1007/s40520-018-0894-0).