Accepted final version of the article:

García-Herranz, S., Díaz-Mardomingo, M. d. C., Suárez-Falcón, J. C., Rodríguez-Fernández, R., Peraita, H., & Venero, C. (2022). Normative data for the Spanish version of the California Verbal Learning Test (TAVEC) from older adults. *Psychological Assessment*, 34(1), 91-97.

DOI: https://doi.org/10.1037/pas0001070

ABSTRACT

The aim of this study was to establish normative data for the Spanish version of the California Verbal Learning test, the Test de Aprendizaje Verbal España-Complutense (TAVEC). Through different sub-tests, the TAVEC allows verbal learning and episodic memory to be evaluated, an assessment that was carried out on a sample of 382 cognitive healthy Spanish individuals aged 60–90 years old. Unlike the participant's educational level, their age and sex significantly influenced performance in the TAVEC. We provide tables that allow the scaled scores obtained with this test to be adjusted for age and other tables with the relevant adjustments for sex. The normative data obtained in this study will help more precisely interpret the performance of older Spanish adults in the TAVEC, enhancing the utility of this neuropsychological test to evaluate verbal learning and episodic memory in clinical settings, and in relation to healthy aging.

Keywords: Spanish version of the California Verbal Learning Test; TAVEC; normative data; aging; episodic verbal memory.

Public Significance Statements

The present study provides updated normative data for the Spanish version of the California Verbal Learning test, the Test de Aprendizaje Verbal España-Complutense (TAVEC) in Spanish older adults. The norms produced here are age and sex adjusted, enhancing the utility of this neuropsychological test to evaluate verbal learning and episodic memory in both clinical and research contexts.

INTRODUCTION

As aging is the main risk factor for cognitive decline and dementia (Hebert et al., 2010), there is a clear need for standardized cognitive assessment tests that are sufficiently sensitive to identify age-related cognitive decline at its earliest stages. Using such tools will enable appropriate psychological and pharmacological interventions to be designed and implemented. Establishing normative data for cognitive tests in older adults is also essential to screen for impairment and to track the progression of cognitive decline (O'Connell et al., 2019). In this sense, the use of relatively large samples is necessary to ensure that the scores obtained are unbiased estimates of the population parameters (Miller et al., 2015). Normative data enable an individual's specific scores to be compared with their reference group based on the variables that may be associated with cognitive performance, such as age, sex and educational level (Busch & Chapin, 2008). In Spain, the need to develop normative data for cognitive aging has been satisfied through different studies, among which the Spanish Multicenter Normative Studies (NEURONORMA Project) (Peña-Casanova et al., 2009a) stands out, providing normative data for a set of cognitive tests (Adrián et al., 2015; Contador et al., 2016; Peña-Casanova et al., 2009b).

Age-related episodic memory failure is one of the earliest cognitive signs of Mild Cognitive Impairment (MCI), a condition associated with a higher risk of progression to dementia (Albert et al., 2011; García-Herranz et al., 2016; Petersen et al., 2001). Several cognitive tools are now available to assess verbal learning and memory, such as the California Verbal Learning Test (CVLT) (Delis et al., 1987), the Rey Auditory Verbal Learning Test (RAVLT) (Rey, 1964), the Buschke Selective Reminding test (BSRT) (Buschke, 1973), and the free and cued selective reminding test (FCSRT) (Buschke, 1984). The Spanish version of the CVLT, the Test de Aprendizaje Verbal España-Complutense (TAVEC) (Benedet &

Alejandre, 1998, 2014), is one of the cognitive tests most widely used with Spanish speaking populations and it has some advantages over other verbal memory tests. First, it has good psychometric properties, with a higher reliability coefficient than those observed in the aforementioned verbal learning tests (Benedet & Alejandre, 1998). Second, it is a sensitive tool to classify amnestic MCI subtypes -amnestic and multidomain- in both clinical and research contexts (Juncos-Rabadán et al., 2012; Rodríguez-Rodríguez et al., 2008; Sales-Galán, 2013). Third, the structure and design of this test allows the three basic processes involved in memory to be distinguished: encoding, storage, and retrieval of verbal information. Indeed, the different sub-tests assess the capacity to learn auditory material, short and long-delay recall, the use of semantic cues and the recognition memory of previously presented information. Finally, this test has ecological validity as it uses lists of words like "shopping lists".

Although normative data that is adapted to Spanish populations exists for the TAVEC, giving it added value as a diagnostic test, this dates back to 1998. However, it must be borne in mind that the sociodemographic characteristics of older Spanish people, as well as their cultural and social influences, have changed significantly over the past 23 years. Moreover, of the 1,015 individuals initially sampled, only 175 were over 65 years of age, such that older individuals were apparently underrepresented at that time. In addition, among the older population the normative data for the Spanish TAVEC only discriminated between 3 age groups: 55-64; 65-74; and >74 years old. It is well known that verbal episodic memory declines with advancing age, an effect that is believed to begin between the ages of 65 and 70 years old (Messinis et al., 2016; Kramer et al., 2020; van der Elst et al., 2005). Therefore, it would appear to be necessary to elaborate more detailed and accurate age-specific normative data to define whether an individual´s performance is within the expected range for normal aging, or whether it can be considered as memory impairment.

Another important limitation of the existing normative data for the TAVEC concerns the absence of any adjustment for the effects of sex or educational level, since the z-scores are only age-corrected. Indeed, significant sex-differences have been reported in verbal memory tests, with females consistently outperforming males (Gale et al., 2007; Kramer et al., 2020). Regarding the influence of formal education on the performance of different verbal episodic memory tests, the data from different normative studies are controversial or inconsistent. Some reported a significant influence of education (Peña-Casanova et al., 2009b; Stricker et al., 2020), whereby individuals with a higher educational level perform better, while others found no such effect (Ivnik et al., 1997; Kramer et al., 2020; Messinis et al., 2016).

The quality of the normative data for cognitive assessment is crucial to maximize the diagnostic and descriptive accuracy of subjects' test scores. Despite the fact that a new edition of the TAVEC was produced in 2014 and a parallel version of this test was published (Nieto et al., 2014), the associated normative data were not revised. Given that the scores associated with any normative data must be re-evaluated periodically (Evers et al., 2010), the main objective of this work was to provide updated normative data for the TAVEC derived from a large sample of healthy, monolingual Spanish older adults, evaluating the effects of age, sex and education in order to adequately stratify the normative data obtained.

METHODS

Participants

We selected 457 healthy, Caucasian, Spanish, monolingual (Spanish-speaking) older adults who were already enrolled on a larger, ongoing longitudinal research study to determine the prevalence and progression of MCI in the Autonomous Community of Madrid (Spain) (García-Herranz et al., 2016; García-Herranz et al., 2019). The participants were selected

according to the following criteria: a) they were between 60 and 90 years old; b) they were living independently; c) they were monolingual Spanish-speaking and; d) scored ≥25 on the Mini-Mental State Examination (MMSE - Spanish version MEC-35) (Lobo et al., 1979). They were recruited through an advertisement placed in several cultural and educational centers for older adults in urban and suburban areas. The study protocol was approved by the ethical committee of the Universidad Nacional de Educación a Distancia (UNED, Spain) and it was carried out in accordance with the tenets of the Helsinki Declaration governing research on human subjects. The exclusion criteria applied were: (a) previously diagnosed MCI or a neurodegenerative disease; (b) a disabling chronic disease; (c) a psychiatric disorder (e.g., major depression); (d) a marked neurological abnormality; (e) severe sensory deficit; (f) diabetes; (g) having experienced a prior cerebrovascular accident; (h) traumatic brain injury; (i) metabolic disease; and (k) the use of any drug that may affect cognitive function (i.e.: anticholinergic drugs, corticoids, benzodiazepines, etc.). Similarly, participants were excluded if the score with the shortened Spanish version of the Geriatric Depression Scale (GDS-15) (Sheikh & Yesavage, 1986) was ≥6 or if the Blessed Dementia Scale (BDS) - Part A (Blessed et al., 1968) returned a value >3.5. Accordingly, the final study cohort consisted of 382 research participants after 75 participants were excluded.

Instruments

Like the original versions of the CVLT (Delis et al., 1987), the Spanish version of the CVLT – the Test de Aprendizaje Verbal España-Complutense (TAVEC) (Benedet & Alejandre, 1998, 2014) – is composed of two lists, A and B, each with 16 items from four semantic categories. List B of TAVEC has 16 words that differ from those in list A. Both lists contain words that belong to four semantic categories (4 words for each category). Two of the semantic categories (fruits and spices) are common to both lists, while the other two are specific to each list. List A is presented first, followed by a free recall session. List A is

presented over five learning trials in which the subject recalls as many items as possible after each trial. After Trial 5, each participant is presented with an interference list (List B). Subsequently, the participant is again asked to recall the items from List A (short delay free recall (SDFR)) and then cued with each of the four categories of words in List A (short delay cued recall (SDCR)). After a 20-minute delay, a long delay free recall (LDFR) and long delay cued recall (LDCR) trials are administered. The last part of the test is the recognition (R) task, in which a list of 44 items is read to the participants and they are asked to indicate whether or not each item was on List A.

Of all the test variables, we focused on the following 9 variables: 1) performance in trial 1, as a measure of immediate memory; 2) performance in trial 5, as a measure of immediate memory and learning; 3) the total number of words recalled over the five trials, as an indicator of the subjects' learning curve from List A; 4) performance with List B to assess retroactive interference; 5) SDFR; 6) SDCR; 7) LDFR; 8) LDCR; and 9) recognition (R) hits.

Procedure

All the research participants were interviewed in order to collect their personal and sociodemographic data, and all the neuropsychological assessments were performed by psychologists with adequate training. The participants all provided their written consent before they were examined using a neuropsychological battery, in accordance with the guidelines of the UNED.

Statistical analysis

All the statistical analyses were carried out using the IBM SPSS Statistics 25 package. We applied the same uniform normative procedure as that used previously in the MOANS studies (Ivnik et al., 1992; Ivnik et al., 1997) and in the Spanish NEURONORMA studies (Peña-Casanova et al., 2009a). The uniform normative procedure followed was:

1) An overlapping interval strategy (Pauker, 1988) was adopted to maximize the number of participants that contribute to the normative distribution at each mid-point age interval. The age distribution of the cohort allowed us to calculate norms for 7 mid-point age range groups: 60-66, 67-69, 70-72, 73-75, 76-78, 79-81 and 82-90 (see supplementary Table S1).

2) The correlation coefficients (r) and coefficients of determination (R^2) were calculated to analyze the effects of sociodemographic variables (age, years of education and sex) in all the TAVEC sub-tests.

3) We created tables of age-adjusted normative values. The frequency distribution of the raw scores was converted into age-adjusted scaled scores (SS_A) to ensure a normal distribution (Ivnik et al., 1992; Peña-Casanova et al., 2009a). A cumulative frequency distribution of the raw scores was generated for each age range so that percentile ranks could be assigned to direct scores based on their place within a distribution. We then transformed the direct scores to SS_A scores (i.e.: scaled scores from 2 to 18) based on the percentiles. This transformation produces a normalization of the distribution (mean = 10 and SD = 3) that makes it possible to perform a linear regression analysis.

4) By convention, when determination coefficients (R^2) values of univariate regressions are less than .05 (i.e.: 5% of the shared variance), the variable studied does not have a relevant effect and it would not be necessary to control for this (Lucas et al., 2005). Such was the case of the variable years of formal education, since in our sample none of the R^2 values in

the TAVEC sub-tests exceeded .05. Therefore, we first applied a correction for age and subsequently, for sex using the equation: $SS_{A\&S} = SS_A - (\beta * Sex)$

RESULTS

Descriptive statistics for all the sociodemographic variables are provided in Table 1. The participant's ages ranged between 60 and 90 years old (mean age = 68.98 years, SD = 5.78) and their level of education ranged from having no formal studies to having post-graduate degrees (mean education = 13 years, SD = 5.98) (see Table 1). For selected participants, the mean score in MMSE Spanish version was 32.8 (range from 25 to 35), indicating no impairment in cognitive functioning.

(Insert Table 1 here)

There was a significant and negative correlation between age and education in the sample (r = -.22, p <.01). By contrast, there were no differences between male and female participants in terms of age, although there were differences in terms of the years of education (t = 5.84, p<.001, mean education of males = 15.07 years, SD = 6.02; mean education of females = 11.58 years, SD = 5.53). In our sample, the Cronbach's alpha coefficient for both the semantic categories and for the total of the items in the five trials was 0.88 and 0.91, respectively. These coefficients are higher than those obtained from the original scores from the TAVEC (0.80 and 0.86, respectively), as well as those from the CVLT study (0.74 and 0.69 respectively).

Raw scores from the TAVEC are presented from the whole sample, and they are also presented segregated by sex in Table 2.

(Insert Table 2 here)

Correlations (Pearson's, r) and shared variance (R^2) of the TAVEC sub-tests scores with age (years), education (years) and sex were assessed (Table 3). Accordingly, age and sex, but not education, accounted for the raw score variance of all measures. Age-adjusted scaled scores (SS_A) for the TAVEC sub-tests are presented in the supplementary material (Tables S2–8).

(Insert Table 3 here)

The regression coefficients (β) for each of the TAVEC scores and the adjustment of these scores according to sex were calculated (see supplementary Table S9 and S10, respectively), bearing in mind that the value obtained in the $SS_{A\&S}$ is rounded down to the nearest integer using the TRUNC (variable) function (for example, 10.75 becomes 10).

Discussion

In the present study we have obtained updated normative data for one of the most relevant and widely used cognitive tests used to assess Spanish speaking older adults, the TAVEC. We assessed the influence of sociodemographic variables like age, years of education and sex in a large sample of cognitively healthy Spanish speakers in the age range of 60 to 90 years old. A decline in performance was detected in all the immediate and delayed recall sub-tests with increasing age, indicating the importance of using age stratified TAVEC normative data. In addition, we observed a significant effect of sex in all the TAVEC sub-tests explored, evidence of the need to adjust these scores for sex. No significant effect of the years of education was observed in our sample.

We initially investigated the existence of significant effects of age on the performance in the TAVEC, detecting a decline in all the TAVEC sub-test scores with advancing age. These findings are in line with previous normative studies in which age influenced verbal learning and episodic memory performance (Kramer et al., 2020; Thielen et al., 2019; van der Elst et

al., 2005). Furthermore, the original TAVEC study also found an age-related decline in all the TAVEC measures but, compared to our sample, participants older than 74 years showed lower raw scores in the TAVEC sub-tests suggesting that verbal memory decline is slower in our sample. It should be noted that the original TAVEC was based on the study of 1,015 adults aged from 16 years old, of which only 175 were over 65 years of age (Benedet & Alejandre, 1998). These older participants were only differentiated into two age groups (normative data for 65-74 years and for older than 74 years) regardless of sex. However, here we included 382 cognitive healthy adults, 227 females and 155 males, in the age range of 60 to 90 years old, with an even distribution across seven age groups. Therefore, our older adult sample is more than twice the size of that in the original TAVEC study.

In addition, we examined whether the years of formal education influenced performance in the TAVEC sub-tests in our older adult cohort. Our regression analysis indicated that the years of education explained less than 5% of the variance in the TAVEC scores. Therefore, the level of formal education did not affect the normative TAVEC data, as reported in the original version (Benedet & Alejandre, 1998), even though our sample has a higher proportion of participants with secondary and further education than in the original TAVEC study. There is a considerable debate concerning the influence of education on the performance of verbal episodic memory tests and there are some inconsistent findings in the literature. Thus, a significant contribution of this variable has been seen in some studies (Peña-Casanova et al., 2009b; Stricker et al., 2020; Thielen et al., 2019), suggesting some benefits of education on learning capacity (Labos et al., 2008), whereas formal education did not apparently influence this type of memory elsewhere (Ivnik et al., 1997; Kramer et al., 2020; Messinis et al., 2016; Speer et al., 2014). In this regard, it has been suggested that education does not explain more than the level of intelligence could (Strauss et al., 2006).

We also evaluated the potential effect of sex on the performance in the TAVEC sub-tests, which revealed that females performed better than males in terms of total recall, in both the short and long-term free recall tests, as well as in the tests of recall with cues (both short and long-term). Our findings are also consistent with other normative data studies on verbal episodic memory tests, indicating the existence of sex differences and pointing to a better performance of females than males (Kramer et al., 2020; Messinis et al., 2016; Stricker et al., 2020; Sundermann et al., 2019; van der Elst et al., 2005). Sex differences in the TAVEC performance may be related to distinct approaches in the coding and learning of verbal information, as well as with the type of organizational strategy used (Delis et al., 1988; Gale et al., 2007).

While we provide age and sex adjusted norms here, it should be noted that the previous TAVEC norms were not adjusted by sex. The data obtained here support the need to provide sex-adjusted normative data for a number of reasons, not least because we found significant sex differences in the performance of the TAVEC. Previous normative data studies recommended that sex differences should be taken into account in the evaluation of age-related changes in cognitive function, especially when verbal learning and episodic memory is assessed (Gale et al., 2007; Stricker et al., 2020; van der Elst et al., 2005). Indeed, sex-adjusted normative data will increase sensitivity for amnestic MCI diagnosis and therefore, special attention should be paid to the fact that if normative data adjusted for sex are not used, the diagnosis of MCI in females could be delayed until a more advanced disease state is reached than in males (Stricker et al., 2020; Sundermann et al., 2019).

Another important difference with respect to the normative data originally proposed by Benedet & Alejandre (1998) can be found in the presentation. Instead of proposing z-scores based on means and standard deviation, the revised normative data presented here were obtained by applying the statistical procedures for data analyses previously used to establish

the study's norms (Ivnik et al., 1992; Ivnik et al., 1997; Peña-Casanova et al., 2009a). Thus, we included scaled scores ($SS_{A\&S}$) following age and sex adjustment. This type of presentation permits a more precise interpretation of the individual's memory performance, and it enables clinicians and researchers to make reliable comparisons using a variety of commonly used neuropsychological tests. The normative data produced here will enable memory deficits to be detected early and will improve diagnostic accuracy in both sexes, reducing the risk of misdiagnosis, especially in cases of suspected MCI or of any other neurodegenerative pathology like AD (Peña-Casanova et al., 2009b; Speer et al., 2014; Stricker et al., 2020; Sundermann et al., 2019). As MCI is usually defined by memory scores at least 1.5 SD below the expected levels in the absence of dementia, and in conjunction with subjective memory complaints (Petersen et al., 2001), in our normative data adjusted score values ≤ 5 would correspond to a performance below normal that might indicate a possible memory deficit.

The present work has some limitations, one of which is related to the selection of the participants. We are aware of the limited representation of very old participants (aged between 75 and 90 years) and the smaller number of males. Alternatively, it must be borne in mind that the normative data presented here was obtained for a monolingual Spanish population drawn from urban and suburban areas, and therefore it may not be possible to generalize this to Spanish speaking countries other than Spain. Future studies should also include bilingual participants, as well as participants from rural areas, both for comparison and to enhance the generalizability of the results. In addition, longitudinal studies may help to develop robust normative standards for TAVEC, reflecting only the most cognitively stable sample of older adults.

In conclusion, we believe that the normative data presented here further supports the use of the TAVEC as one of the reference tests to assess verbal learning and memory in monolingual Spanish older populations, both in a clinical and research context.

Conflict of interests

The authors have no potential conflicts of interests to declare.

REFERENCES

Adrián, J.A., Jorquera, J., & Cuetos, F. (2015). NEUROBEL: Breve batería neuropsicológica de evaluación del lenguaje oral en adultos-mayores. Datos normativos iniciales. *Revista de Logopedia, Foniatría y Audiología, 35,* 101-113.

Albert, M.S., DeKosky, S.T., Dickson, D., Dubois, B., Feldman, H.H., Fox, N.C., ... & Phelps, C.H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, *7*, 270-279. https://doi.org/10.1016/j.jalz.2011.03.008

Benedet, M.J., & Alejandre, M.A. (1998). TAVEC. Test de Aprendizaje Verbal España-Complutense. TEA Ediciones.

Benedet, M.J., & Alejandre, M.A. (2014). TAVEC. Test de Aprendizaje Verbal España-Complutense, 2ª edición revisada. TEA Ediciones.

Blessed, G., Tomlinson, B.E., & Roth, M. (1968). The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *The British Journal of Psychiatry: 114*, 797–811. https://doi.org/10.1192/bjp.114.512.797

Busch, R.M., & Chapin, J.S. (2008). Review of normative data for common screening measures used to evaluate cognitive functioning in elderly individuals. *The Clinical Neuropsychologist, 22,* 620–650. https://10.1080/13854040701448793

Buschke, H. (1973). Selective reminding analysis of memory and learning. *Journal of Verbal Learning and Verbal Behavior*, *12*, 543-550.

Buschke, H. (1984). Cued recall in amnesia. *Journal of Clinical Neuropsychology, 6*, 433-440.

Contador, I., Almondes, K., Fernández-Calvo, B., Boycheva, E., Puertas-Martín, V., Benito-León, J., & Bermejo-Pareja, F. (2016). Semantic verbal fluency: normative data in older Spanish adults from NEDICES population-based cohort. *Archives of Clinical Neuropsychology, 31,* 9. https://doi:10.1093/arclin/acw071

Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (1987). CVLT, California Verbal Learning Test: Adult Version: Manual. Psychological Corporation.

Delis, D.C., Freeland, J., Kramer, J.H., & Kaplan, E. (1988). Integrating clinical assessment with cognitive neuroscience: construct validation of the California Verbal Learning Test. *Journal of Consulting and Clinical Psychology, 56,* 123. https://doi.org/10.1037/0022-006X.56.1.123

Evers, A., Sijtsma, K., Lucassen, W., & Meijer, R.R. (2010). The Dutch Review Process for Evaluating the Quality of Psychological Tests: History, Procedure, and Results. *International Journal of Testing, 10*, 295-317. https://doi.org/10.1080/15305058.2010.518325

Gale, S.D., Baxter, L., Connor, D.J., Herring, A., & Comer, J. (2007). Sex differences on the Rey Auditory Verbal learning test and the brief visuospatial memory test–revised in the elderly: Normative data in 172 participants. *Journal of Clinical and Experimental Neuropsychology, 29*, 561-567. https://doi.org/10.1080/13803390600864760

García-Herranz, S., Díaz-Mardomingo, M.C., & Peraita, H. (2016). Neuropsychological predictors of conversion to probable Alzheimer disease in elderly with mild cognitive impairment. *Journal of Neuropsychology, 10*, 239–255. https://doi.org/10.1111/jnp.12067

García-Herranz, S., Díaz-Mardomingo, M. C., Venero, C., & Peraita, H. (2019). Accuracy of verbal fluency tests in the discrimination of mild cognitive impairment and probable Alzheimer's disease in older Spanish monolingual individuals. *Aging, Neuropsychology, and Cognition*, *27*(6), 826-840. https://doi.org/10.1080/13825585.2019.1698710

Hebert, L.E., Bienias, J.L., Aggarwal, N.T., Wilson, R.S., Bennett, D.A., Shah, R.C., & Evans, D.A. (2010). Change in risk of Alzheimer disease over time. *Neurology*, *75*, 786-791. https://doi.org/10.1212/WNL.0b013e3181f0754f

Ivnik, R.J., Malec, J.F., Smith, G.E., Tangalos, E.G., Petersen, R.C., Kokmen, E., & Kurland, L.T. (1992). Mayo's older Americans normative studies: updated AVLT norms for ages 56 to 97. *The Clinical Neuropsychologist*, *6*, 83-104. https://doi.org/10.1076/jcen.24.2.214.995

Ivnik, R. J., Smith, G. E., Lucas, J. A., Tangalos, E. G., Kokmen, E., & Petersen, R. C. (1997). Free and cued selective reminding test: MOANS norms. *Journal of Clinical and Experimental Neuropsychology, 19*, 676-691. https://doi.org/10.1080/01688639708403753.

Juncos-Rabadan, O., Pereiro, A.X., Facal, D., Rodriguez, N., Lojo, C., Caamaño, J.A., ... & Eiroa, P. (2012). Prevalence and Correlates of Cognitive Impairment in Adults with Subjective Memory Complaints in Primary Care Centers. *Dementia and Geriatric Cognition Disorders*, *33*, 226–232. https://doi.org/10.1159/000338607

Kramer, A.O., Casaletto, K.B., Umlauf, A., Staffaroni, A.M., Fox, E., You, M., & Kramer, J.H. (2020). Robust normative standards for the California Verbal Learning Test (CVLT) ages 60–89: A tool for early detection of memory impairment. *The Clinical Neuropsychologist*, *34*, 384-405. https://10.1080/13854046.2019.1619838

Labos, E., Trojanowski, S., & Ruiz, C. (2008). Prueba de recuerdo libre/facilitado con recuerdo inmediato. *Revista Neurológica Argentina, 33*, 50-66.

Lobo, A., Ezquerra, J., Gómez Burgada, F., Sala, J.M., & Seva Díaz, A. (1979). El Mini Examen Cognoscitivo (un "test" sencillo, práctico, para detectar alteraciones intelectivas en pacientes médicos). *Actas Luso-Españolas de Neurología, Psiquiatría y Ciencias Afines, 3*,189-202.

Lucas, J.A., Ivnik, R.J., Smith, G.E., Ferman, T.J., Willis, F.B., Petersen, R.C., & Graff-Radford, N.R. (2005). Mayo's Older African Americans Normative Studies: Norms for the Boston naming test, controlled oral word association, category fluency, animal naming, token test, WRAT-3 reading, Trail making test, Stroop test, and judgement of line orientation. *The Clinical Neuropsychologist, 19*, 243–269. https://doi.org/10.1080/13854040590945337

Messinis, L., Nasios, G., Mougias, A., Politis, A., Zampakis, P., Tsiamaki, E., ... & Papathanasopoulos, P. (2016). Age and education adjusted normative data and discriminative validity for Rey's Auditory Verbal Learning Test in the elderly Greek population. *Journal of Clinical and Experimental Neuropsychology*, *38*, 23-39. https://doi.org/10.1080/13803395.2015.1085496

Miller, I.N., Himali, J.J., Beiser, A.S., Murabito, J.M., Seshadri, S., Wolf, P.A., & Au, R. (2015). Normative Data for the Cognitively-Intact Oldest-Old: The Framingham Heart Study. *Experimental Aging Research, 41*, 386–409. http://doi.org/10.1080/0361073X.2015.1053755.

Nieto, A., Hernández-Rodríguez, E., Hernández-Torres, A., Velasco Rodríguez-Solis, P., Hess-Medler, S., Machado-Fernández, A., Molina-Rodríguez, Y., & Barroso, J. (2014). Versión paralela del test de aprendizaje verbal España-Complutense (TAVEC). *Revista de Neurología, 58*, 95. https://doi.org/10.33588/rn.5809.2013568

O'Connell, M.E., Tuokko, H., Kadlec, H., Griffith, L.E., Simard, M., Taler, V., ... & Raina, P. (2019). Normative comparison standards for measures of cognition in the Canadian Longitudinal Study on Aging (CLSA): Does applying sample weights make a difference? *Psychological Assessment, 31*, 1081–1091. https://doi.org/10.1037/pas0000730

Pauker J.D. (1988). Constructing overlapping cell tables to maximize the clinical usefulness of normative test data: rationale and an example from neuropsychology. *Journal of Clinical Psychology*, *44*, 930-933. https://doi.org/10.1002/1097-4679(198811)44:6<930::AID-JCLP2270440613>3.0.CO;2-H

Peña-Casanova, J., Blesa, R., Aguilar, M., Gramunt-Fombuena, N., Gómez-Ansón, B., Oliva, R., ... & Sol, J.M. (2009a). Spanish multicenter normative studies (NEURONORMA project): Methods and sample characteristics. *Archives of Clinical Neuropsychology*, *24*, 307-319. https://doi.org/10.1093/arclin/acp027

Peña-Casanova, J., Gramunt-Fombuena, N., Quiñones-Úbeda, S., Sánchez-Benavides, G., Aguilar, M., Badenes, D., ... & NEURONORMA Study Team. (2009b). 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*, 371-393. https://doi.org/10.1093/arclin/acp041

Petersen, R.C., Doody, R., Kurz, A., Mohs, R.C., Morris, J.C., Rabins, P.V., ... & Winblad, B. (2001). Current concepts in mild cognitive impairment. *Archives of Neurology, 58*, 1985-1992. https://doi:10.1001/archneur.58.12.1985

Rey, A. (1964). L'examen Clinique en Psychologie Paris: Presses Universitaires de France.

Rodríguez-Rodríguez N., Juncos-Rabadán O., & Facal D. (2008). Cognitive markers to discriminate between mild cognitive impairment and normal ageing. *Revista Española de Geriatría y Gerontología*, *43*(5), 291-298. https://10.1016/s0211-139x(08)73571-6

Sales-Galán, A., Meléndez-Moral, J.C., & Mayordomo-Rodríguez, T. (2013). Using a cognitive plasticity measure to detect mild cognitive impairment. *Archives of Clinical Neuropsychology*, *28*(8), 763-70. https://doi.org/10.1093/arclin/act064.

Sheikh, J.I., & Yesavage, J.A. (1986). Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Clinical Gerontologist*,5, 165–73.

Speer, P., Wersching, H., Bruchmann, S., Bracht, D., Stehling, C., Thielsch, M., ... & Lohmann, H. (2014). Age-and gender-adjusted normative data for the German version of Rey's Auditory Verbal Learning Test from healthy subjects aged between 50 and 70 years. *Journal of Clinical and Experimental Neuropsychology*, 36, 32-42. https://doi.org/10.1080/13803395.2013.863834

Strauss, E., Sherman, E. M., & Spreen, O. (2006). A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary. New York: Oxford University Press.

Stricker, N.H., Christianson, T.J., Lundt, E.S., Alden, E.C., Machulda, M.M., Fields, J.A., ... & Petersen, R.C. (2020). Mayo Normative Studies: Regression-Based Normative Data for the Auditory Verbal Learning Test for Ages 30–91 Years and the Importance of Adjusting for Sex. *Journal of the International Neuropsychological Society*, 1-16. https://doi.org/10.1017/S1355617720000752

Sundermann, E.E., Maki, P., Biegon, A., Lipton, R.B., Mielke, M.M., Machulda, M., ... & Alzheimer's Disease Neuroimaging Initiative. (2019). Sex-specific norms for verbal memory tests may improve diagnostic accuracy of amnestic MCI. *Neurology, 93*, e1881-e1889. https://doi.org/10.1212/WNL.0000000000008467

Thielen, H., Verleysen, G., Huybrechts, S., Lafosse, C., & Gillebert, C.R. (2019). Flemish normative data for the Buschke selective reminding test. *Psychologica Belgica, 59,* 58. http://doi.org/10.5334/pb.486

van der Elst, W., Van Boxtel, M.P., Van Breukelen, G.J., & Jolles, J. (2005). Rey's verbal learning test: normative data for 1855 healthy participants aged 24-81 years and the influence of age, sex, education, and mode of presentation. *Journal of the International Neuropsychological Society, 11*, 290–302. https://doi.org/10.1017/S1355617705050344

Table 1. *Sociodemographic characteristics of the sample (n = 382)*

	N	%
Sex		
Female	227	59.4
Male	155	40.6
Age (years)		
60-64	93	24.3
65-69	138	36.1
70-74	89	23.3
75-79	43	11.3
80+	19	5.0
Education (years)		
≤5	33	8.6
6-7	36	9.4
8-9	45	11.8
10-11	53	13.9
12-13	36	9.4
14-15	42	11.0
≥16	137	35.9

Table 2. *Means and standard deviation for all the measures of the TAVEC for the total sample and for males and females*

Measures	Total N = 382		Female N =227		Male N =155		*p-value*	*Effect Size*
	Mean	SD	Mean	SD	Mean	SD		
Trial 1	6.01	2.04	6.38	2.12	5.46	1.79	.000	.046
Trial 5	11.33	2.69	12.03	2.37	10.30	2.80	.000	.066
List A	46.81	11.06	49.90	10.48	42.29	10.32	.000	.067
List B	5.53	2.04	5.87	2.07	5.02	1.90	.000	.042
Short delay free recall	9.05	3.26	9.75	3.22	8.02	3.04	.000	.068
Short delay cued recall	10.23	2.75	10.88	2.68	9.28	2.59	.000	.081
Long delay free recall	9.65	3.37	10.42	3.29	8.53	3.19	.000	.076
Long delay cued recall	10.38	2.94	11.03	2.88	9.43	2.78	.000	.072
Recognition	14.34	1.74	14.52	1.76	14.07	1.68	.014	.016

Statistically significant $p \leq .05$

Table 3. *Correlation (r) and shared variances (R^2) of the raw scores with age, years of education and sex*

Raw score	Age (years)		Education		Sex	
TEST	*r*	R^2	*r*	R^2	*r*	R^2
Trial 1	-0.18 [b]	0.03	0.06	0.00	0.22 [b]	0.05 [c]
Trial 5	-0.22 [b]	0.05 [c]	0.03	0.00	0.31 [b]	0.10 [c]
List A	-0.22 [b]	0.05 [c]	0.03	0.00	0.34 [b]	0.11 [c]
List B	-0.19 [b]	0.04	0.11	0.01	0.20 [b]	0.04
Short delay free recall	-0.18 [b]	0.03	0.05	0.00	0.26 [b]	0.07 [c]
Short delay cued recall	-0.16 [b]	0.03	0.06	0.00	0.28 [b]	0.08 [c]
Long delay free recall	-0.21 [b]	0.04	0.03	0.00	0.27 [b]	0.08 [c]
Long delay cued recall	-0.16 [b]	0.03	0.04	0.00	0.27 [b]	0.07 [c]
Recognition	-0.11 [a]	0.01	0.03	0.00	0.13 [a]	0.02

[a] Significant correlation at $p \leq .01$ (2-tails); [b] Significant correlation at $p \leq .05$ level (2-tails); [c] $R^2 \geq .05$

Supplementary material

Legend

To use the tables correctly, select the appropriate table that corresponds to the age of the individual, find their SSA, and then refer to the corresponding $SSA_{\&}S$ to correct for sex.

Supplementary Table S1. *TAVEC mid-point groups (according to Pauker's strategy)*

Groups	Midpoint age	Age range for midpoint	Age range for the norms	Sample size
1	65	≤66 (60-66)	60-70	250
2	68	67-69	63-73	258
3	71	70-72	66-76	223
4	74	73-75	69-79	156
5	77	76-78	72-82	96
6	80	79-81	75-85	57
7	83	≥82 (82-90)	≥78 (78-90)	35

Supplementary Table S2. *Age-adjusted TAVEC scores (SS_A) for the age range 60-66 (age range for norms = 60–70)*

Percentil range	Scaled score	TAVEC								
		TRIAL 1	TRIAL 5	LIST A	LIST B	SDFR	SDCR	LDFR	LDCR	R
<1	2	<2	<5	<23	<2	<1	<4	<0	<3	<10
1	3	2	--	23	--	1	4	0	3	10
2-3	4	--	5	24-27	--	2-3	5	1-3	4	--
4-6	5	--	6	28-29	2	4	6	4	5-6	11
7-12	6	3	7-8	30-35	3	5	--	5	--	12
13-20	7	4	9	36-40	--	6	7	6	7	--
21-30	8	--	10	41-43	4	7	8	7-8	8-9	13
31-42	9	5	11	44-46	--	8	9	9	10	--
43-57	10	--	--	47-50	5	9	10	10	--	14
58-69	11	6	12	51-53	6	10	11	11	11	--
70-79	12	7	13	54-57	--	11	12	12	12	15
80-87	13	8	--	58-61	7	12	13	13	13	--
88-93	14	9	14	62-63	8	13	14	14	14	--
94-96	15	10	15	64-66	9	14	--	--	--	--
97-98	16	--	--	67	10	15	15	15	15	--
99	17	11	16	68-70	11	16	16	16	16	16
>99	18	>11	>16	>70	>11	>16	>16	>16	>16	>16
	Sample size 250									

SDFR, Short delay free recall; SDCR, Short delay cued recall; SDFR, Long delay free recall; LDCR, Long delay cued recall; R, Recognition.

Supplementary Table S3. *Age-adjusted TAVEC scores (SS_A) for the age range 67-69 (age range for norms = 63–73)*

		TAVEC								
Percentile range	Scaled score	TRIAL 1	TRIAL 5	LIST A	LIST B	SDFR	SDCR	LDFR	LDCR	R
<1	2	<2	<4	<21	<1	<1	<4	<0	<2	<10
1	3	--	4	21-22	1	1	4	0	2	10
2-3	4	2	5	23-26	--	2-3	5	1-3	3-4	--
4-6	5	--	6	27-29	2	--	--	4	5	11
7-12	6	3	7-8	30-33	--	4-5	6	5	6	12
13-20	7	--	9	34-39	3	6	7	6	7	--
21-30	8	4	--	40-41	4	7	8	7	8	13
31-42	9	5	10	42-45	--	8	9	8-9	9	--
43-57	10	--	11	46-49	5	9	10	10	10	14
58-69	11	6	12	50-53	--	10	11	11	11	--
70-79	12	7	13	54-57	6	11	12	12	12	15
80-87	13	8	--	58-60	7	12	--	13	13	--
88-93	14	9	14	61-63	8	13	13	14	14	--
94-96	15	10	15	64-66	9	14	14	--	15	--
97-98	16	--	--	67-68	10	15	15	15	--	--
99	17	11	16	69-71	11	16	16	16	16	16
>99	18	>11	>16	>71	>11	>16	>16	>16	>16	>16
	Sample size 258									

SDFR, Short delay free recall; SDCR, Short delay cued recall; SDFR, Long delay free recall; LDCR, Long delay cued recall; R, Recognition.

Supplementary Table S4. *Age-adjusted TAVEC scores (SS_A) for the age range 70-72 (age range for norms = 66–76)*

Percentile range	Scaled score	TAVEC								
		TRIAL 1	TRIAL 5	LIST A	LIST B	SDFR	SDCR	LDFR	LDCR	R
<1	2	<2	<4	<22	<0	<0	<4	<0	<2	<10
1	3	--	4	22-23	0	0	4	0	2	10
2-3	4	2	5	24-26	1	1-2	5	1-3	3-4	--
4-6	5	--	6	27-29	--	3	--	4	5	11
7-12	6	3	7	30-34	2	4-5	6	5	6	12
13-20	7	--	8	35-38	3	6	7	6	7	--
21-30	8	4	9	39-41	--	7	8	7	8	13
31-42	9	--	10	42-44	4	8	9	8	9	--
43-57	10	5	11	45-47	5	9	10	9	10	14
58-69	11	6	12	48-52	--	10	11	10-11	11	15
70-79	12	7	13	53-56	6	11	--	12	12	--
80-87	13	--	--	57-59	7	12	12-13	13	13	--
88-93	14	8	14	60-62	--	13	--	--	14	--
94-96	15	9	15	63-66	8	14	14	14	--	--
97-98	16	10	--	67	9	15	15	15	15	--
99	17	11	16	68-70	10	16	16	16	16	16
>99	18	>11	>16	>70	>10	>16	>16	>16	>16	>16
	Sample size 223									

SDFR, Short delay free recall; SDCR, Short delay cued recall; SDFR, Long delay free recall; LDCR, Long delay cued recall; R, Recognition.

Supplementary Table S5. *Age-adjusted TAVEC scores (SS_A) for the age range 73-75 (age range for norms = 69–79)*

Percentile range	TAVEC									
	Scaled score	TRIAL 1	TRIAL 5	LIST A	LIST B	SDFR	SDCR	LDFR	LDCR	R
<1	2	<2	<4	<20	<0	<1	<4	<0	<3	<7
1	3	--	4	20-23	0	--	4	--	3	7
2-3	4	2	5	24-28	1	1	--	0	4	8-10
4-6	5	--	6	29	--	2-3	5	1-3	--	11
7-12	6	3	7	30-33	2	4	6	4	5-6	12
13-20	7	--	8	34-36	3	5	--	5	7	--
21-30	8	4	9	37-40	--	6	7-8	6	--	13
31-42	9	--	10	41-42	4	7	--	7-8	8-9	--
43-57	10	5	11	43-46	5	8	9	9	10	14
58-69	11	6	--	47-51	--	9	10	10	11	15
70-79	12	--	12	52-55	6	10	11	11	--	--
80-87	13	7	13	56-58	--	11	12	12	12	--
88-93	14	8	14	59-61	7	12-13	13	13	13	--
94-96	15	9	15	62-65	8-9	14	14	14	14	--
97-98	16	10	--	66-69	--	--	15	15	--	--
99	17	11	16	70-71	10	15	16	16	15	16
>99	18	>11	>16	>71	>10	>15	>16	>16	>15	>16
	Sample size 156									

SDFR, Short delay free recall; SDCR, Short delay cued recall; SDFR, Long delay free recall; LDCR, Long delay cued recall; R, Recognition.

Supplementary Table S6. *Age-adjusted TAVEC scores (SS_A) for the age range 76-78 (age range for norms = 72–82)*

		TAVEC								
Percentile range	Scaled score	TRIAL 1	TRIAL 5	LIST A	LIST B	SDFR	SDCR	LDFR	LDCR	R
<1	2	<2	<3	<19	<0	<0	<3	<0	<2	<6
1	3	--	3	19	0	0	3	--	2	6
2-3	4	--	4-5	20-28	1	1-2	4	0-1	3-4	7-9
4-6	5	2	6	29-30	--	--	5	2-3	--	10
7-12	6	3	7	31-33	2	3	--	4	5	11
13-20	7	--	8	34-35	3	4-5	6	5	6	12
21-30	8	4	9	36-38	--	6	7-8	6	7-8	13
31-42	9	--	--	39-42	4	7	--	7	9	14
43-57	10	5	10	43-45	--	8	9	8-9	10	--
58-69	11	6	11	46-49	5	9-10	10-11	10	11	15
70-79	12	--	12	50-56	--	--	--	11	12	--
80-87	13	7	13	57-58	6	11	12-13	12	--	--
88-93	14	8	14	59-61	7	12-13	14	13	13-14	--
94-96	15	9	15	62-65	8	14	--	14	--	--
97-98	16	10	16	66-70	9-10	15	15-16	15-16	15-16	16
99	17	--	--	--	--	--	--	--	--	--
>99	18	>10	>16	>70	>10	>15	>16	>16	>16	>16
	Sample size 96									

SDFR, Short delay free recall; SDCR, Short delay cued recall; SDFR, Long delay free recall; LDCR, Long delay cued recall; R, Recognition.

Supplementary Table S7. *Age-adjusted TAVEC scores (SS_A) for the age range 79-81 (age range for norms = 75–85)*

		TAVEC								
Percentile range	Scaled score	TRIAL 1	TRIAL 5	LIST A	LIST B	SDFR	SDCR	LDFR	LDCR	R
<1	2	<1	<2	<12	<2	<0	<3	<0	<4	<5
1	3	--	--	12	--	--	--	--	--	--
2-3	4	--	2-4	13-18	--	--	3	--	4	5-6
4-6	5	1	5	19-27	--	0-1	4	0-1	--	7-8
7-12	6	2	6-7	28-32	2	2	5	2-3	5	9-11
13-20	7	3	--	33-34	--	3	6	4	6	12
21-30	8	--	8-9	35-37	3	4-5	7	5	7	13
31-42	9	4	--	38-40	4	6-7	8	6	8	--
43-57	10	--	10	41-44	--	8	9	7-8	9-10	14
58-69	11	5	11	45-47	5	9	10	9-10	11	--
70-79	12	6	12	48-54	--	10	11	11	12	15
80-87	13	7	--	55-57	6	11	12-13	12	--	--
88-93	14	8	13	58-60	--	12-13	--	13	13	--
94-96	15	9	14	61	7-8	14	14	14	14	--
97-98	16	10	15-16	62-66	9	15	15	15-16	15-16	16
99	17	--	--	--	--	--	--	--	--	--
>99	18	>10	>16	>66	>9	>15	>15	>16	>16	>16
	Sample size 57									

SDFR, Short delay free recall; SDCR, Short delay cued recall; SDFR, Long delay free recall; LDCR, Long delay cued recall; R, Recognition.

Supplementary Table S8. *Age-adjusted TAVEC (SS_A) for the age range 82-90 (age range for norms = 78–90)*

Percentile range	Scaled score	TAVEC								
		TRIAL 1	TRIAL 5	LIST A	LIST B	SDFR	SDCR	LDFR	LDCR	R
<1	2	<1	<2	<12	<2	<0	<4	<1	<2	<5
1	3	--	--	--	--	--	--	--	--	--
2-3	4	--	--	12-13	--	0	--	--	2	--
4-6	5	1	2-3	14-19	--	1-2	--	1	3-4	5-7
7-12	6	2	4-5	20-25	--	--	4	2-3	5	8-9
13-20	7	3	6-7	26-32	2	3	5	4	--	10-11
21-30	8	--	--	33-34	3	4	6-7	5	6	12
31-42	9	4	8-9	35-37	4	5-6	8	6	7-8	13
43-57	10	--	10	38-42	--	7-8	9	7-8	9	14
58-69	11	5	11	43-46	5	9	10	9-10	10	--
70-79	12	6	12	47-48	--	10	11	11	11	15
80-87	13	7	--	49-54	6	11	12	--	12	--
88-93	14	8-9	--	55-60	--	12	13-14	12-13	13-14	--
94-96	15	--	13	--	7-8	13-14	--	--	15	--
97-98	16	10	14	61	9	15	15	14	16	16
99	17	--	--	--	--	--	--	--	--	--
>99	18	>10	>14	>61	>9	>15	>15	>14	>16	>16
	Sample size 35									

SDFR, Short delay free recall; SDCR, Short delay cued recall; SDFR, Long delay free recall; LDCR, Long delay cued recall; R, Recognition.

Supplementary Table S9. *Regression coefficients to correct for the effect of sex (β)*

TAVEC	β
Trial 1	1.234276
Trial 5	2.034077
List A	2.160523
List B	1.241495
Short delay free recall	1.694118
Short delay cued recall	1.786585
Long delay free recall	1.725082
Long delay cued recall	1.723177
Recognition	1.338016

Supplementary Table 10. *TAVEC sex-adjusted scores (SS_{AS}= SS_A - (β * Sex) (Male = 0 and Female = 1).*

	Trial 1		Trial 5		List A		List B		SDFR		SDCR		LDFR		LDCR		R	
SS_A	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F	M	F
2	2	0	2	0	2	0	2	0	2	0	2	0	2	0	2	0	2	0
3	3	1	3	0	3	0	3	1	3	1	3	1	3	1	3	1	3	1
4	4	2	4	1	4	1	4	2	4	2	4	2	4	2	4	2	4	2
5	5	3	5	2	5	2	5	3	5	3	5	3	5	3	5	3	5	3
6	6	4	6	3	6	3	6	4	6	4	6	4	6	4	6	4	6	4
7	7	5	7	4	7	4	7	5	7	5	7	5	7	5	7	5	7	5
8	8	6	8	5	8	5	8	6	8	6	8	6	8	6	8	6	8	6
9	9	7	9	6	9	6	9	7	9	7	9	7	9	7	9	7	9	7
10	10	8	10	7	10	7	10	8	10	8	10	8	10	8	10	8	10	8
11	11	9	11	8	11	8	11	9	11	9	11	9	11	9	11	9	11	9
12	12	10	12	9	12	9	12	10	12	10	12	10	12	10	12	10	12	10
13	13	11	13	10	13	10	13	11	13	11	13	11	13	11	13	11	13	11
14	14	12	14	11	14	11	14	12	14	12	14	12	14	12	14	12	14	12
15	15	13	15	12	15	12	15	13	15	13	15	13	15	13	15	13	15	13
16	16	14	16	13	16	13	16	14	16	14	16	14	16	14	16	14	16	14
17	17	15	17	14	17	14	17	15	17	15	17	15	17	15	17	15	17	15
18	18	16	18	15	18	15	18	16	18	16	18	16	18	16	18	16	18	16

M, male; F, female; SDFR, Short delay free recall; SDCR, Short delay cued recall; SDFR, Long delay free recall; LDCR, Long delay cued recall; R, Recognition.