





## Research Article

# Phonological fluency norms for Spanish middle-aged and older adults provided by the SCAND initiative (P, M, & R)

M.L. Delgado-Losada<sup>1</sup>, S. Rubio-Valdehita<sup>2</sup> , R. López-Higes<sup>1</sup>, M. Campos-Magdaleno<sup>3</sup> , M. Ávila-Villanueva<sup>4</sup>, B. Frades-Payo<sup>4</sup> and C. Lojo-Seoane<sup>3</sup>

<sup>1</sup>Department of Experimental Psychology, Complutense University of Madrid, 28223 Madrid, Spain, <sup>2</sup>Department of Social Work and Differential Psychology, Complutense University of Madrid, 28223 Madrid, Spain, <sup>3</sup>Department of Developmental and Educational Psychology, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain and <sup>4</sup>CIEN Foundation, Carlos III Institute of Health, 28029, Madrid, Spain

### Abstract

**Objective:** Verbal fluency tests are quick and easy to administer neuropsychological measures and are regularly used in neuropsychological assessment. Additionally, phonological fluency is a widely used paradigm that is sensitive to cognitive impairment. This paper offers normative data of phonological verbal fluency (letters P, M, R) for Spanish middle- and older-aged adults, considering sociodemographic factors, and different measures such as the total number of words, errors (perseveration and intrusions), and 15 sec-segmented scores. **Method:** A total of 1165 cognitively unimpaired participants aged between 50 and 89 years old, participated in the study. Data for P were obtained for all participants. Letters M and R were also administered to a subsample of participants (852) aged 60 to 89 years. In addition, errors and words produced every 15 seconds were collected in the subsample. To verify the effect of sociodemographic variables, linear regression was used. Adjustments were calculated for variables that explained at least 5% of the variance ( $R^2 \geq .05$ ). **Results:** Means and standard deviations by age, scaled scores, and percentiles for all tests across different measures are shown. No determination coefficients equal to or greater than .05 were found for sex or age. The need to establish adjustments for the educational level was only found in some of the measures. **Conclusions:** The current norms provide clinically useful data to evaluate Spanish-speaking natives from Spain aged from 50 to 89 years. Specific patterns of cognitive impairment can be analyzed using these normative data and may be important in neuropsychological assessment.

**Keywords:** normative data; phonological verbal fluency; ageing; SCAND; Spain; descriptive survey study

(Received 25 October 2022; final revision 13 April 2023; accepted 18 April 2023)

### Introduction

Verbal fluency (VF) involves expressive language abilities, storage of language knowledge, and executive functions. VF tests are typically brief assessment instruments that permit the evaluation of these cognitive processes with simple administration and scoring procedures (Lezak et al., 2012; Lezak et al., 2012; Pekkala, 2012; Strauss et al., 2006). Due to their high sensitivity to neurological damage, they are widely used in clinical evaluation and research in various areas such as neuropsychology, speech therapy, linguistics, and medicine (see for example: Catani et al., 2013; Faroqi-Shah & Milman, 2018; Herbert et al., 2014).

Both in clinical practice and in research, local normative data are essential for the correct interpretation of the data obtained by a person in one or several neuropsychological test/s, no matter whether the objective is diagnosis or elaboration of a judgment about their cognitive state. In the clinical setting, a patient is compared with a group of people who have the same sociodemographic characteristics that he/she has. If the main purpose is this and it is necessary to choose some reference standards, the first ones that should be rejected

are those that have been developed with the smallest sample size (Mitrushina et al., 2005). Those that were developed perhaps more than a decade ago, or that may be biased by cultural differences, or by other factors (low educational level, etc.) should be rejected too. Normative studies provide what has been called "clinical comparison data" (Mitrushina et al., 2005), which represents the range of performance on a test of different groups characterized by medical, psychiatric, and/or neurological criteria, who present homogeneous demographic traits.

Previous initiatives have provided norms for phonological fluency (PF) tests in different populations of Spanish-speaking natives from Spain: Peña-Casanova et al., (2009) for community-dwelling cognitively normal adults ( $n = 346$ ), ranging from 50 to 94 years of age, used the letters P, M, and R; Casals-Coll et al. (2013) for a younger population, adults between 18 and 49 years old ( $n = 179$ ), used the letters P, M, and R; and Lubrini et al. (2022) that assessed participants from 17 to 100 years old ( $n = 257$ ), utilized the letters F, A, and S. If we consider what has been mentioned before about the choice of reference normative

**Corresponding author:** Susana Rubio-Valdehita; Email: [srubiova@ucm.es](mailto:srubiova@ucm.es)

**Cite this article:** Delgado-Losada M.L., Rubio-Valdehita S., López-Higes R., Campos-Magdaleno M., Ávila-Villanueva M., Frades-Payo B., & Lojo-Seoane C. Phonological fluency norms for Spanish middle-aged and older adults provided by the SCAND initiative (P, M, & R). *Journal of the International Neuropsychological Society*, 1–11, <https://doi.org/10.1017/S1355617723000309>

Copyright © INS. Published by Cambridge University Press 2023. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

data, it is evident that the studies carried out with a small sample (Casal-Coll *et al.*, and Lubrini *et al.*) would be the first to discard. Considering the date of publication, the work by Peña-Casanova *et al.*, is 14 years old, so it could be considered outdated. Furthermore, it is important to note that in Peña-Casanova *et al.*, study, they used an overlapping strategy to distribute their sample ( $n = 346$ ) across 10 groups aged from 55 to more than 80 years, thus artificially increasing the number of cases in each age range. It is evident that updated norms with larger samples are necessary for both Spanish young and older adult populations.

According to the statistical data of the National Institute of Statistics (INE) of Spain, elderly people now represent approximately 19.5% of the total population. The mean age of the population stands at 43.81 years when in 1970 it was 32.7. According to the INE projection, in 2035 there could be more than 12.8 million older people, 26.5% of the total population (Pérez *et al.*, 2022). The need for valid and accurate normative data is especially important for older people, as this group is at special risk of cognitive impairment or dementia. As a result, it is necessary to develop global programs and increase resources focused on promoting prevention and early diagnosis. In this context, normative studies conducted by the SCAND initiative ([www.scandcognition.org](http://www.scandcognition.org)), such as the one presented here for PF with a sample of middle-aged and older adults, make sense.

Due to its importance, it is necessary to mention other precedents in the international sphere. For example, the Neuropsychological Norms for the US-Mexico Border Region in Spanish (NP-NUMBRS) project (Rivera Mindt *et al.*, 2021), and specifically, for its relation to the present study, the normative data published on VF (Marquine *et al.*, 2021). It should also be noted the Mayo Clinic initiative that started in the 90s (Ivnik *et al.*, 1996) and has provided norms for the most important neuropsychological test in different populations (see for example: Lucas *et al.*, 2005).

Numerous studies have pointed out the influence of socio-demographic variables as possible moderators of VF, with age, gender, and education being the most studied (Henry & Phillips, 2006; López-Higes *et al.*, 2022; Mathuranath *et al.*, 2003). Some authors have reported that men perform worse than women in PF tasks (Loonstra *et al.*, 2001), whereas many studies have not shown such differences (Costa *et al.*, 2014; Khalil, 2010; Kozora & Cullum, 1995; Peña-Casanova *et al.*, 2009; Tombaugh *et al.*, 1999). Some studies have found the different effects of age when comparing semantic fluency (SF) and PF (Santos Nogueira *et al.*, 2016; Tombaugh *et al.*, 1999), whereas others did not (Khalil, 2010; Loonstra *et al.*, 2001). However, the impact of the educational level on the performance of VF tasks has been widely recognized in neuropsychological research (Lubrini *et al.*, 2022; Peña-Casanova *et al.*, 2009). It has been found that higher education levels are associated with the production of more words (Oberg & Ramírez, 2006). The significant effect of educational level on PF could be related to the fact that these tasks are more demanding than semantic tasks and more sensitive to executive dysfunction (Shores *et al.*, 2006). Formal education may increase vocabulary and consequently greater verbal lexical retrieval capacity (Henry & Phillips, 2006). In fact, education is the major factor that contributes to the performance in PF.

The performance on a PF test is usually evaluated by the total number of correct words given within the time limit (Lezak *et al.*, 2012; Pekkala, 2012; Strauss *et al.*, 2006). However, this score provides little information about the cognitive processes underlying fluency performance, thus some authors have proposed

additional measures, such as the error types in PF, which typically include perseverations (repetition of the same correct word) and intrusions (words with another initial letter) (Thiele *et al.*, 2016). Although errors are relatively rare in normative data, investigating the number and types of errors is useful in research and clinical practice as it is not only the decline in response counts that indicates pathology, but also alterations in performance patterns.

Research suggests that different mechanisms probably underlie perseverations and intrusions. Perseverations have been linked to a frontal lobe dysfunction characterized by intellectual rigidity and inability to shift mental sets (Miller & Cohen, 2001) which is common in neurologic disorders. By recording the number of perseveration errors, one might have more information about the status of the central executive component of working memory.

Intrusions on a PF task could be executive errors (e.g., forgetting the rules, losing the set, and using rules from a different fluency trial; McDowd *et al.*, 2011) or phonetic/spelling related (a word with a similar sound but incorrect letter; Rofes *et al.*, 2019). These errors have been studied in different pathologies such as brain injury, Parkinson's disease, or Alzheimer's disease (Smith *et al.*, 2020), suggesting that intrusions are typical of senile dementia of the Alzheimer's type and may help distinguish it from other causes of dementia. Regarding the clinical usefulness of the error patterns analysis, it is important to remember that it is better to interpret them within the context of the whole neuropsychological assessment.

In some studies, time has also been considered. In these cases, the number of words generated during four segments of 15 seconds each has been evaluated. In general, participants produce most of the words in the early stages of the task (first 30 seconds) using a semi-automatic rapid retrieval process. Cognitive demand is not uniform throughout the PF task since, as time progresses, lexical retrieval becomes more difficult, and therefore fewer words are produced in the final moments of the task (Venegas & Mansur, 2011). Furthermore, it has been found that the educational level of the participants has a positive effect on the first-time segments in both PF and SF, while age was not significant (Venegas & Mansur, 2011).

Several studies in languages other than English have aimed to adapt their letter sets so they might pose a similar difficulty to "F, A, and S" in English. For instance, the "P, M, and R" set has been proposed as more appropriate to be used with Spanish speakers for several reasons (Fortuny *et al.*, 1998). First, words beginning with "F" are rare in Spanish. Although words beginning with "A" are common, the starting "HA" (as in "hábito", habit) is also very frequent, which may be disadvantageous for people with low levels of literacy since the letter H is silent in Spanish. Finally, the "S" sound may be confusing in some regions of Spain and Latin America ("C" in sequences "CE"/"CI", -as in "celebración", celebration-, and "Z" in sequences "ZA/ZO/ZU", -as in "zapato", shoe-, are pronounced like "S"), which again poses a disadvantage for people with low literacy.

The availability of normative data on PF, adjusted for age, education, or sex, can help in the early detection of cognitive impairment and the measurement of clinically significant changes.

The main purpose of this paper is to offer updated normative data of PF (P, M, R, and P + M + R) for Spanish-speaking middle-aged and older adults natives from Spain (over 50 years old), considering sociodemographic factors (age, education, and sex), which have been elaborated with a large sample. Instead of providing only normative data for a single outcome measure (the total of words evoked) as other previous studies did, a novel contribution of the present work is the inclusion of both errors and

word production during 15-second segments as additional measures for analyses.

## Method

### Participants

Participants were selected using the following inclusion criteria: (1) community-dwelling individuals; (2) over 50 years of age; (3) Mini-Mental State Exam (MMSE; Lobo et al., 1999) greater or equal to 24 points; (4) Geriatric Depression Scale of 15 items version (GDS-15; Yesavage et al., 1983) below or equal to 9 points; (5) normal cognitive development, not meeting diagnostic criteria for Mild Cognitive Impairment (MCI) (Petersen, 2004) in at least two previous consecutive assessments; (6) being able to manage an independent life without any severe mental disorder (cognitive or psychiatric) impeding daily functioning; (7) normal or corrected hearing and vision; (8) basic reading comprehension and writing abilities in Spanish; and (9) signed written informed consent.

A total of 1165 Spanish speakers without cognitive impairment aged between 50 and 89 years were recruited. In the study, the sampling was non-probabilistic incidental, most of the participants lived in urban areas (93%) and did not receive any type of remuneration for their participation. All participants were asked to produce words that began with P. Letters M and R were also administered to a subsample of participants (852) aged 60 to 89 years.

Before unifying data from the full 'P' sample with the 'M and R' subsample, we verified that there were no statistically significant differences in performance that could be due to the group [ $F(1,1165) = 2.44, p = .118$ ] or its interaction with socio-demographic variables: group  $\times$  sex [ $F(1,1165) = 2.35, p = .125$ ]; group  $\times$  educational level [ $F(3,1163) = 1.97, p = .117$ ]; group  $\times$  age group [ $F(2,1164) = 1.33, p = .264$ ]; group  $\times$  sex  $\times$  educational level [ $F(3,1163) = 2.19, p = .087$ ]; group  $\times$  sex  $\times$  age group [ $F(2,1164) = 0.35, p = .703$ ]; group  $\times$  educational level  $\times$  age group [ $F(6,1160) = 0.80, p = .565$ ]; or group  $\times$  sex  $\times$  educational level  $\times$  age group [ $F(5,1161) = 0.86, p = .503$ ].

Table 1 shows the distribution by sex and educational level for the total sample and the subsample. Regarding Spanish people aged 50 or more, they completed their studies under the educational law enacted in 1970 or under the law that precedes it, established in 1953, that was reformed in 1967. In respect of university studies, there was a law in 1943 substituted later in 1970 for another that distinguished different levels (diplomado, licenciado, doctor, trad. English: diplomate, graduate, doctorate). To cope with the heterogeneity of levels or grades (EGB, elemental bachelor, superior bachelor, BUP, etc.) we used the following equivalents, related to international categories. We have considered educational level as an ordinal variable (with values ranging from 0 to 3). That is, 'Without formal education' (0): less than 6 years of schooling; 'Primary studies' (1): between 6 and 11 years of schooling; 'Secondary studies' (2): between 12 and 15 years of schooling; 'Higher studies' (3): more than 15 years of formal education. The level "Without formal education" corresponded to people who are literate but could only go to school for 2 or 3 years (they only have basic reading and writing skills and simple mathematical calculation). The mean age of the total sample was 72.08 years ( $SD = 6.46$ ) and of the subsample was 74.01 years ( $SD = 3.83$ ). The median of educational level was 2 for the total sample and the subsample. Sex was coded as 0 for females and 1 for males. The percentage of women was 65% in the total sample and 63% in the subsample.

**Table 1.** Distribution of women and men by educational level and age group

	Total sample			Subsample		
	Female	Male	Total	Female	Male	Total
Education						
Without formal education	136	46	182	104	36	140
Primary	250	91	341	167	64	231
Secondary	187	112	299	139	92	231
Higher	186	157	343	125	125	250
Age						
50–59	64	11	75			
60–69	117	57	174	43	22	65
70–79	508	295	803	438	260	698
80–89	70	43	113	54	35	89

Note. Educational level is an ordinal variable, with values ranging from 0 to 3. 'Without formal education' (0): less than 6 years of schooling; 'Primary studies' (1): between 6 and 11 years of schooling; 'Secondary studies' (2): between 12 and 15 years of schooling; 'Higher studies' (3): more than 15 years of formal education

The Spanish Consortium for Ageing Normative Data (SCAND) initiative takes the data from three Spanish cohort studies, Aging Brain Projects of the Complutense University of Madrid, the Vallecas Project, and the Compostela Aging Study. The SCAND initiative was developed with the aim of sharing data on the Spanish middle-aged and old adult population provided that the above-mentioned studies share a set of neuropsychological tests in their evaluation protocols. All participants in the total sample were selected given that they had been involved in different studies about the aging process and were recruited between 2008 and 2019.

All studies complied with the ethical standards of the Declaration of Helsinki and were approved by the local Ethics Committees of the participant institutions.

### Instruments

In the present study, three letters were considered (P, M, and R), all of which were part of the comprehensive neuropsychological assessment protocol used by each of the SCAND research groups, which included screening tests, scales, and other tests belonging to different cognitive domains (memory, executive functions, and language). Initially, the PF task with the letter P was included in the protocol. Although this is common among a significant number of clinicians, we also found some cases in which the letters M or R were used in isolation. For this reason, we later began to use the three letters (P, M, and R) with participants. Participants were asked to generate as many words as possible beginning with these initial letters in 60 seconds. For each letter, the number of correct words was registered, excluding intrusions and perseverations. Errors (perseverations and intrusions) and words produced every 15 seconds were recorded only for the subsample.

### Procedure

All participants completed a structured interview to collect sociodemographic data, screening tests, and an extensive neuropsychological assessment, including memory, executive functions, and language tests, administered by neuropsychologists well-trained in the use of neuropsychological assessment tools. All participants were informed about the main research aspects, and they signed a written informed consent before performing any study procedure.

According to standard instructions (e.g., Lezak et al., 2012), participants were asked to generate in 60 seconds as many words as possible that began with each letter (presented always in this order

and consecutively: P, M, and R), excluding proper names and repetitions of the same word with different endings. Participants started naming words beginning with each letter and then two neuropsychologists recorded correct words, perseverations, and intrusions in the order that they were generated. Two raters were used to ensure the reliability of the scoring procedure. The inter-rater reliability was near .99 in all cases.

### Data analyses

The statistical procedure was as follows: First, the cumulative frequency distribution of the raw scores was generated. Percentile ranges were assigned to the raw scores depending on their place in distribution. Then the percentile ranges were converted to scaled scores (*ss*) (range 2 to 19) using the formula  $ss = 10 + 3 * Z$  where *Z* is the normalized standard score corresponding to the percentile. This transformation of raw scores produced a normal distribution ( $M = 10$  and  $SD = 3$ ), which allows the application of linear regressions to test the effect of sociodemographic variables and to calculate the scaled-adjusted scores (*ssadjusted*).

Secondly, the effects of age, education, and sex were verified. For each letter, three univariate regressions were calculated on *ss* with age, education, and sex as predictors. Corrections were only applied for those sociodemographic variables that yielded a significant regression coefficient ( $p < .05$ ) and that also explained more than 5% of the variance (Lee, 2014). Finally, adjustments were made according to age, education, and sex on the *SS*, using the following formula:

$$ssadjusted = ss - [B_1 * (Age - Mean_1) + B_2 * (Education - Median_2) + B_3 * (Sex - Mode_3)]$$

In the formula, and due to the level of measurement of each of the variables, the mean is subtracted from age (continuous ratio level), the median from education (ordinal level), and the mode from sex (nominal level), so that the adjusted scores provide a better-standardized reference. All analyses were performed with SPSS version 25.

### Results

Table 1 showed the distribution of women and men by educational level and age group.

Correlations between the individual letter scores and each one with the P + M + R score were high and statistically significant ( $r_{P-M} = .703$ ;  $r_{P-R} = .718$ ;  $r_{M-R} = .747$ ;  $r_{P-PMR} = .896$ ;  $r_{M-PMR} = .902$ ;  $r_{R-PMR} = .910$ ;  $p \leq .001$  in all cases).

### Norms for the total number of correct words produced by participants

Table 2 shows the descriptive statistics of the total number of correct words by sociodemographic variables. Age was categorized only to show its distribution in the tables but was included as continuous in all statistical analyses.

Determination coefficients corresponding to education exceed the criterion in P, M, R, and P + M + R score, explaining 19.4% ( $r = .444$ ;  $p < .001$ ;  $R^2 = .194$ ), 22.9% ( $r = .478$ ;  $p < .001$ ;  $R^2 = .229$ ), 21.8% ( $r = .467$ ;  $p < .001$ ;  $R^2 = .218$ ), and 26.6% ( $r = .516$ ;  $p < .001$ ;  $R^2 = .266$ ), of the variance, respectively, however, age and sex did not reach the 5% criterion regarding the percentage of variance explained ( $R^2 < .05$  in all cases).

Table 3 includes unadjusted scaled scores (*ss*) and percentile ranges corresponding to the total number of words evoked by participants.

The adjustments by education in P, M, R, and P + M + R indicated that for people without formal education +2 points must be added to their *ss* score for P, M and R, and +3 points when P + M + R score is considered. For individuals with primary education, +1 must be added for P, M, R or P + M + R, and when the educational level is higher, the *ssadjusted* score for P, M, R, or P + M + R must be -2 points. No adjustments are needed for the secondary education level. To explain how to use tables to select the correct scaled score given a raw score and how to use the correction on unadjusted scaled scores if needed, let us consider an example. If a patient without formal education produced 16 words with the letter P, we first locate the raw score in Table 3, then we see the percentile range at the left (69-75) and the corresponding unadjusted scaled score ( $ss = 12$ ). To adjust the unadjusted scaled score according to her/his educational level "without formal education" it is necessary to correct his/her unadjusted scaled score by adding two points ( $ssadjusted = 14$ ). For a patient with the same raw score but with a higher educational level, their unadjusted scaled score should be corrected by subtracting two points ( $ssadjusted = 10$ ).

### Norms for errors (perseverations and intrusions)

Table 4 summarizes descriptive statistics of perseverations and intrusions by sociodemographic variables. Data for errors were obtained only from the subsample of 852 participants.

Scaled scores and percentile ranges corresponding to perseveration and intrusion errors across letters are shown in Table 5. None of the sociodemographic variables could explain at least 5% of the variance, so no adjustments are required ( $R^2 < .05$  in all cases).

### Norms for the total number of words produced every 15 seconds

Descriptive statistics corresponding to the total number of words produced every 15 seconds are shown in Table 6.

Table 7 shows unadjusted scaled scores (*ss*) and percentile ranges corresponding to this measure across letters.

The only variable that explained at least 5% of the variance was education (**From 0 to 15 sec.:**  $r = .299$ ,  $p = .001$ ,  $R^2 = .089$  for P;  $r = .379$ ,  $p = .001$ ,  $R^2 = .143$  for M;  $r = .347$ ,  $p = .001$ ,  $R^2 = .121$  for R; and  $r = .417$ ,  $p < .001$ ,  $R^2 = .174$  for P + M + R. **From 16 to 30 sec.:**  $r = .338$ ,  $p = .001$ ,  $R^2 = .114$  for P;  $r = .348$ ,  $p = .001$ ,  $R^2 = .121$  for M;  $r = .373$ ,  $p = .001$ ,  $R^2 = .139$  for R; and  $r = .456$ ,  $p < .001$ ,  $R^2 = .207$  for P + M + R. **From 31 to 45 sec.:**  $r = .345$ ,  $p = .001$ ,  $R^2 = .119$  for P;  $r = .340$ ,  $p = .001$ ,  $R^2 = .116$  for M;  $r = .323$ ,  $p = .001$ ,  $R^2 = .104$  for R; and  $r = .420$ ,  $p < .001$ ,  $R^2 = .177$  for P + M + R. **From 46 to 60 sec.:**  $r = .281$ ,  $p = .001$ ,  $R^2 = .079$  for P;  $r = .306$ ,  $p = .001$ ,  $R^2 = .094$  for M;  $r = .321$ ,  $p = .001$ ,  $R^2 = .104$  for R; and  $r = .424$ ,  $p < .001$ ,  $R^2 = .180$  for P + M + R). Adjustments were only needed for education so that when considering the letter P, *ssadjusted* is  $ss + 1$  for those without formal education and  $ss - 1$  for those with higher education, in all 15-second segments. When the letter is M, in the first segment (0 to 15 sec.) *ssadjusted* is  $ss + 2$  for patients without formal education,  $ss + 1$  for those with primary education, and  $SS - 2$  for those with higher education, and for the rest of the 15-second segments *ssadjusted* is  $ss + 1$  for those without formal education and  $ss - 1$  for those with higher education. Regarding the letter R, in the second segment (16 to 30 sec.)

**Table 2.** Mean (*M*) and standard deviation (*SD*) of total scores by sociodemographic characteristics

	Total P		Total M		Total R		P+M+R	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Sex								
Male	14.46	4.68	12.08	4.36	12.31	4.41	38.65	11.91
Female	14.05	4.55	11.63	4.16	11.89	4.33	37.29	11.67
Total	14.19	4.60	11.80	4.24	12.05	4.36	37.80	11.77
Education								
Without formal education	11.00	3.41	8.55	3.35	8.69	3.04	27.90	7.98
Primary	12.85	4.20	10.24	3.57	10.60	3.77	33.50	9.94
Secondary	14.73	4.26	12.87	3.88	13.12	4.09	40.64	10.80
Higher	16.75	4.30	14.06	3.88	14.28	4.10	44.68	10.57
Total	14.19	4.60	11.80	4.24	12.05	4.36	37.80	11.77
Age								
50–59	17.09	4.98						
60–69	14.42	4.77	11.55	3.50	12.20	4.41	37.95	11.22
70–79	13.96	4.49	11.90	4.26	12.10	4.38	38.01	11.85
80–89	13.60	4.17	11.18	4.50	11.52	4.19	36.02	11.54
Total	14.19	4.60	11.82	4.24	12.05	4.36	37.80	11.77

Note. Educational level is an ordinal variable, with values ranging from 0 to 3. 'Without formal education' (0): less than 6 years of schooling; 'Primary studies' (1): between 6 and 11 years of schooling; 'Secondary studies' (2): between 12 and 15 years of schooling; 'Higher studies' (3): more than 15 years of formal education.

**Table 3.** Unadjusted scaled scores (*ss*) and percentile ranges corresponding to the total number of words. How to use the table: first locate the raw score, then see the percentile range at the left and the corresponding scaled score. For information about demographic adjustments to these scaled scores see below

<i>ss</i>	Total P		Total M		Total R		P+M+R	
	Percentile ranges	Raw scores	Percentile ranges	Raw scores	Percentile ranges	Raw scores	Percentile ranges	Raw scores
2	≤1	1–3	≤1	1–2	<1	1–2	<1	10–11
3	2	4–5			1	3	1	12–14
4	3	6	2	3	3	4	23	15–17
5	5	7	4–7	4–5	6	5	3–5	18–19
6	10	8	10	6	9	6	6–11	20–23
7	15	9	16	7	15–20	7–8	12–20	24–27
8	22–29	10–11	22	8	30	9	21–27	28–30
9	38	12	31–41	9–10	39	10	28–42	31–34
10	48–56	13–14	49	11	49–56	11–12	43–55	35–38
11	63	15	57–66	12–13	65	13	56–67	39–42
12	69–75	16–17	74–80	14–15	72–78	14–15	68–78	43–47
13	80–86	18–19	86	16	83	16	79–87	48–52
14	90.6	20	90	17	88–92	17–18	88–92	53–55
15	94–95	21–22	94–96	18–19	94	19	93–95	56–58
16	96–98	23–24	97	20	98	20–21	96–98	59–62
17	99	25–26	99	21	99	22	99	63–67
18	>99	27–28	>99	22–26	>99	23–25	>99	68–74

Note: The adjustments by education in P, M, R, and P + M + R indicated that for people without formal education +2 points must be added to their *ss* score for P, M and R, and +3 points when P + M + R score is considered. For individuals with primary education, +1 must be added for P, M, R or P + M + R, and when the educational level is higher, the *ssadjusted* score for P, M, R, or P + M + R must be –2 points. No adjustments are needed for secondary education level.

*ssadjusted* is *ss* +2 for patients without formal education, *ss* + 1 for those with primary education, and *ss*–2 for those with higher education, and for the rest of 15-second segments *ssadjusted* is *ss* + 1 for those without formal education and *ss*–1 for those with higher education. When the sum P + M + R is considered, *ssadjusted* is *ss* + 2 for patients without formal education, *ss* + 1 for those with primary education, and *ss*–1 for those with higher education, regardless of segment.

## Discussion

The present study provides normative data of phonological VF (letters P, M, R) for Spanish middle- and older-aged adults, considering sociodemographic factors, and different measures such as the total numbers of words, errors (perseveration and intrusions), and 15 second-segmented scores.

Regarding the whole sample it is important to note that the percentage of women with a lower educational level was higher than that of men, a fact that reflects the reality of the Spanish population of people over 50 years of age (Pérez et al., 2020). Therefore, the sociodemographic distribution of the participants can be considered representative of the Spanish population, with the sole exception that our group of men has a higher level of education than usual.

All correlations between individual letters and between those and the total P + M + R score are high and similar, which would indicate that a clinician could use a single letter if she/he needs brevity in her/his evaluation protocol instead of P + M + R total score.

Results regarding the total number of words produced by participants point out that education is the most important variable since it explains between 19 and 23% of the total variance

**Table 4.** Mean (*M*) and standard deviation (*SD*) of perseverations and intrusions by sociodemographic characteristics

	PERSEVERATIONS								INTRUSIONS								
	P		M		R		P+M+R		P		M		R		P+M+R		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Sex																	
Male	.62	.90	.54	.83	.61	.82	1.77	1.82	.16	.45	.23	.63	.21	.53	.60	1.09	
Female	.46	.73	.47	.74	.58	.85	1.51	1.58	.11	.38	.15	.44	.13	.39	.38	.74	
Total	.52	.80	.50	.77	.59	.84	1.60	1.68	.13	.41	.18	.52	.16	.45	.47	.89	
Education																	
Without formal education	.35	.68	.34	.55	.37	.74	1.02	1.46	.20	.53	.26	.62	.24	.55	.70	1.17	
Primary	.44	.69	.45	.72	.56	.82	1.43	1.45	.11	.37	.19	.55	.20	.50	.49	.91	
Secondary	.63	.91	.57	.87	.61	.79	1.81	1.73	.11	.35	.18	.52	.12	.36	.40	.79	
Higher	.57	.84	.56	.83	.72	.93	1.88	1.84	.13	.42	.13	.40	.11	.40	.37	.75	
Total	.52	.80	.50	.77	.59	.84	1.60	1.68	.13	.41	.18	.52	.16	.45	.47	.89	
Age																	
60-69	.49	.73	.48	.70	.52	.64	1.51	1.34	.11	.31	.04	.19	.19	.39	.29	.54	
70-79	.52	.82	.49	.77	.60	.84	1.60	1.70	.13	.42	.17	.49	.17	.47	.46	.88	
80-89	.50	.76	.54	.80	.55	.88	1.58	1.60	.12	.40	.28	.72	.11	.34	.52	.98	
Total	.52	.80	.50	.77	.59	.84	1.60	1.68	.13	.41	.18	.52	.16	.45	.46	.99	

Note. Educational level is an ordinal variable, with values ranging from 0 to 3. 'Without formal education' (0): less than 6 years of schooling; 'Primary studies' (1): between 6 and 11 years of schooling; 'Secondary studies' (2): between 12 and 15 years of schooling; 'Higher studies' (3): more than 15 years of formal education.

**Table 5.** Scaled scores (*ss*) and percentile ranges corresponding to perseverations and intrusions

<i>ss</i>	PERSEVERATIONS								INTRUSIONS							
	P		M		R		P+M+R		P		M		R		P+M+R	
	Pc	Rs	Pc	Rs	Pc	Rs	Pc	Rs	Pc	Rs	Pc	Rs	Pc	Rs	Pc	Rs
9							≤31	0								
11	≤62	0	≤64	0	≤58	0	32-58	1								
12							59-76	2								
13					59-87	1	77-87	3							≤70	0
14	63-90	1	65-89	1					≤89	0	≤86	0	≤87	0		
15					88-96	2	88-94	4-5							71-90	1
16	91-97	2	90-97	2			95-97	6			87-98	1	88-98	1	95-96	2
17					97-99	3			90-98	1	99	2			97-98	3
18	98-99	3	≥98	3-4			98	7	99	2	>99	3-4	99	2	99	4
19	>99	4-7			>99	4-5	≥99	8-12	>99	3-4			>99	3-4	>99	5-6

Note that higher scores indicate more perseverations and intrusions and thus worse performance. Pc = Percentile ranges; Rs = Raw scores

in all letters, and 26.6% in the sum P + M + R. The effect of age was also statistically significant, although it was unable to explain 3% of the variance. Sex does not have any significant effect, in accordance with previous studies (Mathuranath et al., 2003).

Results showing the effect of the education level on the total number of words produced are in accordance with all the studies reviewed. Tombaugh et al. (1999) showed a direct influence of education on PF, accounting for 18.6% of the variance. Mathuranath et al. (2003) concluded with similar results, indicating a significant influence of education on PF, in which participants with higher educational levels present better performance than those with fewer years of schooling. Aziz et al. (2017) showed that educational level significantly influences both phonemic and semantic fluency tasks, with higher educational levels being associated with better performance. This result is a constant across studies and it appears even with qualitative characteristics of task performance such as clustering and switching (Pereira et al., 2018)

Another set of studies have shown the significant effects of educational level and age. Dursun et al. (2002) studied the effects of age and total years of education on vocabulary performance in healthy volunteers. They found that education and age were overall

predictors of total scores, but no correlation was found with sex. Peña-Casanova et al. (2009) reported effects of age and education in different letters, but sex was again not significant. A recent study conducted by Marquine et al. (2021) has found a small effect of age and a medium effect of educational level on PF scores, thus showing a similar pattern to that of the present study.

Word production is largely based on verbal skills (vocabulary), memory retrieval and recall, and executive control processes (self-initiation and monitoring to inhibit repetitions and intrusions) (Shao et al., 2014). VF tasks are multidimensional as they rely on other executive functioning skills as well, such as processing speed, cognitive flexibility, working memory, and sustained attention (Diamond, 2013). PF requires search and retrieval strategies dependent on accessing the mental lexicon. A higher educational level entails a larger lexicon and greater verbal lexical retrieval capacity, as well as the use of more efficient information retrieval strategies (Federmeier et al., 2010). Thus, many factors can influence educational attainment. Formal education may increase vocabulary knowledge, a strong predictor of PF performance with age (Henry & Phillips, 2006), and provides contents and procedures frequently included in cognitive testing (Ardila et al.,

**Table 6.** Mean and standard deviation (in brackets) of the total number of words produced every 15 seconds considering sociodemographic variables

	P				M				R				P+M+R			
	0-15	16-30	31-45	45-60	0-15	16-30	31-45	45-60	0-15	16-30	31-45	46-60	0-15	16-30	31-45	46-60
Sex																
Male	5.56 (1.86)	3.61 (1.61)	3.02 (1.52)	2.91 (1.55)	4.50 (1.67)	3.14 (1.48)	2.70 (1.38)	2.52 (1.47)	4.74 (1.67)	3.25 (1.61)	2.75 (1.42)	2.38 (1.46)	14.83 (4.31)	9.99 (3.74)	8.44 (3.36)	7.77 (3.24)
Female	5.54 (1.75)	3.35 (1.63)	2.83 (1.39)	2.70 (1.49)	4.42 (1.64)	3.01 (1.49)	2.51 (1.44)	2.30 (1.35)	4.70 (1.63)	3.05 (1.48)	2.56 (1.43)	2.30 (1.46)	14.65 (4.17)	9.38 (3.58)	7.88 (3.30)	7.29 (3.21)
Total	5.55 (1.79)	3.45 (1.63)	2.90 (1.44)	2.78 (1.52)	4.45 (1.65)	3.06 (1.49)	2.58 (1.42)	2.38 (1.40)	4.71 (1.64)	3.12 (1.53)	2.63 (1.43)	2.33 (1.46)	14.71 (4.22)	9.61 (3.65)	8.08 (3.33)	7.47 (3.23)
Education																
Without formal education	4.64 (1.49)	2.45 (1.32)	2.12 (1.02)	2.10 (1.33)	3.43 (1.50)	2.16 (1.17)	1.80 (1.13)	1.75 (1.20)	3.74 (1.48)	2.17 (1.16)	1.87 (1.21)	1.51 (1.15)	11.79 (3.55)	6.76 (2.39)	5.76 (2.33)	5.32 (2.53)
Primary	5.14 (1.61)	3.13 (1.49)	2.51 (1.38)	2.47 (1.42)	3.97 (1.54)	2.68 (1.34)	2.24 (1.30)	2.00 (1.33)	4.30 (1.49)	2.68 (1.33)	2.32 (1.38)	2.06 (1.29)	13.39 (3.76)	8.46 (3.12)	7.06 (3.02)	6.52 (2.83)
Secondary	5.82 (1.84)	3.65 (1.57)	3.08 (1.47)	2.85 (1.48)	4.81 (1.51)	3.42 (1.54)	2.78 (1.44)	2.62 (1.27)	5.06 (1.57)	3.50 (1.58)	2.85 (1.31)	2.45 (1.38)	15.71 (3.95)	10.57 (3.62)	8.70 (3.14)	7.89 (2.88)
Higher	6.12 (1.76)	4.05 (1.62)	3.48 (1.37)	3.33 (1.52)	5.12 (1.55)	3.59 (1.40)	3.15 (1.39)	2.88 (1.45)	5.32 (1.58)	3.72 (1.46)	3.15 (1.44)	2.92 (1.57)	16.65 (3.90)	11.38 (3.38)	9.76 (3.18)	9.15 (3.26)
Total	5.55 (1.79)	3.45 (1.63)	2.90 (1.44)	2.78 (1.52)	4.45 (1.65)	3.06 (1.49)	2.58 (1.42)	2.38 (1.40)	4.71 (1.64)	3.12 (1.53)	2.63 (1.43)	2.33 (1.46)	14.71 (4.22)	9.61 (3.65)	8.08 (3.33)	7.47 (3.23)
Age																
60-69	5.65 (1.35)	3.57 (1.46)	3.11 (1.28)	2.68 (1.49)	4.48 (1.34)	2.96 (1.25)	2.48 (1.28)	2.22 (1.36)	4.70 (1.29)	3.22 (1.76)	2.52 (1.15)	2.70 (1.38)	14.67 (2.23)	9.85 (3.54)	8.07 (2.63)	7.59 (3.07)
70-79	5.56 (1.81)	3.46 (1.64)	2.90 (1.45)	2.81 (1.52)	4.49 (1.65)	3.06 (1.49)	2.60 (1.43)	2.39 (1.39)	4.74 (1.64)	3.15 (1.53)	2.67 (1.45)	2.33 (1.47)	14.82 (4.24)	9.66 (3.68)	8.14 (3.33)	7.50 (3.22)
80-89	5.40 (1.79)	3.28 (1.56)	2.87 (1.43)	2.63 (1.53)	4.10 (1.73)	3.10 (1.58)	2.49 (1.43)	2.35 (1.46)	4.52 (1.73)	2.90 (1.43)	2.42 (1.32)	2.24 (1.42)	13.95 (4.31)	9.19 (3.44)	7.70 (3.46)	7.16 (3.38)
Total	5.55 (1.79)	3.45 (1.63)	2.90 (1.44)	2.78 (1.52)	4.45 (1.65)	3.06 (1.49)	2.58 (1.42)	2.38 (1.40)	4.71 (1.64)	3.12 (1.53)	2.63 (1.43)	2.33 (1.46)	14.71 (4.22)	9.61 (3.65)	8.08 (3.33)	7.47 (3.23)

Note. Educational level is an ordinal variable, with values ranging from 0 to 3. 'Without formal education' (0): less than 6 years of schooling; 'Primary studies' (1): between 6 and 11 years of schooling; 'Secondary studies' (2): between 12 and 15 years of schooling; 'Higher studies' (3): more than 15 years of formal education

2000). Cognitively stimulating experiences in early life can enhance brain development and impact cognitive ability later in life (Noble et al., 2015). In line with these arguments, some studies have reported that reading level (a proxy related to cognitive reserve) and PF were moderately correlated (Johnson-Selfridge & Zalewski, 2001). The systematic review conducted by Panico et al. (2022) described two studies that found correlations between cognitive reserve (CR) and PF. Moraes et al. (2013) investigated the correlation between CR (expressed by years of formal education and frequency of reading and writing) and scores on VF tasks (phonemic and semantic fluency), among other tests/tasks. Education showed the best predictive value on PF. Roldan-Tapia et al. (2012) reported that a composite index of CR (including education, occupation, and vocabulary knowledge) significantly correlated with scores on PF. In another interesting study, Kraan et al. (2013) concluded that PF in adults was associated with verbal intellectual function and processing speed.

While traditional normative data studies have examined the total number of words generated as a measure of VF performance, there is evidence suggesting that task performance analysis (errors or temporal analysis) provides valuable additional information (Abwender et al., 2001; Pakhomov et al., 2018). As far as we know, the present study is the first attempt to provide normative data on older Spanish adults concerning the number of perseveration and intrusion errors. Our results showed a small statistically significant effect of educational level that failed to explain at least 5% of the variance of the errors. Effects of age or sex were not significant. Previous studies considering the effects of education on older and middle-aged adults' errors were not found. As a tentative hypothesis, the fact that education explains a percentage under 5% criterion might be related to the small range observed in the

number of errors. Ranges in raw scores are certainly small, given that the sample is composed of healthy older adults.

Concerning our measures of word generation performance on 15-second time intervals, we also did not find previous studies including Spanish normative data in similar measures. Our results indicate that the highest word production occurs in the first 30 seconds of the tests, and a descending curve of word production was observed over the 60-second test period. These results are consistent with those found in the adult population; it has been suggested that production of words is maximal during the initial stages of the task (semi-automatic retrieval) as individuals access their long-term memory, which consists of the greatest frequency, easy-to-retrieve words. When this store is exhausted, the individual attempts to retrieve words from a larger pool of words, making the search process more time-consuming and more difficult (Crowe, 1998; Jacobs et al., 2021; Raboutet et al., 2010). That is, as time on task increased the production decreased, as did the word frequency of the items produced. In a study conducted by Demetriou and Holtzer (2017), healthy older adults were fast and efficient at initiating search processes and retrieving words from memory as evidenced by the larger number of words they produced in the first 20 seconds of the task. However, there was a discrepancy between the first-time intervals and the subsequent two intervals' performance which Demetriou and Holtzer explained assuming that people had to monitor and inhibit responses that had already been given from a large set of evoked words.

Educational level effects appeared in all the 15-sec.-intervals across letters (and in the sum) in the present study. Sex and age had a significant effect on some of the 15-second interval measures but were very weak as they failed to explain 2% of the variance. Similarly, Venegas and Mansur (2011) found that participants'





educational level has a positive effect on the three first quartiles in PF, while age was not significant. As suggested by these authors, the first quartile is dedicated to semi-automatic retrieval, while the other quartiles are implicated in planning, adjusting, and monitoring the performance, to guarantee the generation of items and avoidance of repetitions and intrusions. The effect of educational level on word production in the last 15-second interval is also relevant given that at this point the task requires the greater effortful retrieval processes and people with higher educational attainment should show an advantage given their larger lexicon, greater verbal lexical retrieval capacity, and more efficient information retrieval strategies (Demetriou & Holtzer, 2017). Congruent with this line of reasoning is the study of Sauz on et al. (2011) which revealed a knowledge compensation mechanism in older adults' letter fluency productions that only occurred during the second period (31–60 sec.) and was related to vocabulary level.

We would like to point out as a limitation of this research that we did not use epidemiological recruitment methods, and potentially medical and/or psychological conditions that may interfere with cognition and self-reported mood. Neither of these variables was assessed in this study. We did not recruit illiterate participants because they are very unusual in Spain. Although it provides a greater representation of the Spanish population to the sample of participants, the unequal sex distribution of the sample should be added to the list of the study's limitations. Also, the reader must consider that these normative data are not generalizable to Spanish speakers outside of Spain, provided that other sample demographic characteristics must be similar to those of the normative sample; additionally, the administration and scoring procedures of the test used must be matched too (Mitrushina et al., 2005).

### Conclusions and future directions

The present study provides normative data for healthy older people on the PF task for the letters P, M, and R, and considers errors and production by time segments. The influence of education is in line with other previous studies. These data may also be of considerable use for comparisons with other normative studies in Spain and other countries.

There are very few normative studies for VF tasks for people over 60–75 years of age, and even fewer for Spanish speakers, in addition, to our knowledge, this is the first study to present normative data regarding the number of errors (perseverations and intrusions) and the number of words produced in 15-sec. intervals. Normative data for PF in specific populations are a useful resource for clinical and research studies and may aid in the early detection of cognitive impairment, diagnosis, establishing prognosis, planning treatment, and monitoring clinically significant changes.

It is of interest to further investigate whether a different approach to quantifying performance on PF, such as error testing or time segment generation, is related to prevalent MCI or predictive of its incidence.

**Acknowledgements.** We especially thank all the participants for their totally generous collaboration, without which this study would not have been possible.

**Competing Interest.** None.

**Funding.** This work was supported by the Ministry of Science, Innovation and Universities of Spain (PSI2009-14415-C03-01, PSI2012-38375-C03-01, PSI2010-22224-C03-01, PSI2015-68793-C3-1-R and RTI2018-098762-B-C31)

and the Vallecas Project supported by Reina Sofia Foundation, and PILEP+90 sponsored by Fundaci n General de la Universidad de Salamanca (FGUSAL)-CENIE within the Interreg V-A Program, Spain-Portugal, (POCTEP), 2014-2020.

### References

- Abwender, D. A., Swan, J. G., Bowerman, J. T., & Connolly, S. W. (2001). Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods. *Assessment*, 8(3), 323–338. <https://doi.org/10.1177/107319110100800308>
- Ardila, A., Ostrosky-Solis, F., & Mendoza, V. U. (2000). Learning to read is much more than learning to read: A neuropsychologically based reading program. *Journal of the International Neuropsychological Society*, 6(7), 789–801. <https://doi.org/10.1017/S1355617700677068>
- Aziz, A. K., Khater, M. S., Emar, T., Tawfik, H. M., Rasheedy, D., Mohammedin, A. S., Tolba, M. F., El-Gabry, D. A., & Qassem, T. (2017). Effects of age, education, and gender on verbal fluency in healthy adult Arabic-speakers in Egypt. *Applied Neuropsychology: Adult*, 24(4), 331–341, 2016.
- Casals-Coll, M., S nchez-Benavides, G., Quintana, M., Manero, R. M., Rognoni, T., Calvo, L., Mundi, C., Caltagirone, C., & Pe a-Casanova, J. (2013). Estudios normativos espa oles en poblaci n adulta joven (proyecto NEURONORMA j venes): Normas para los test de fluencia verbal. *Neurolog a*, 28(1), 33–40. <https://doi.org/10.1016/j.nrl.2012.02.010>
- Catani, M., Mesulam, M. M., Jakobsen, E., Malik, F., Martersteck, A., Wieneke, C., Thompson, C. K., & Thiebaut de Schotten, M. Dell'Acqua, F. Weintraub, S. Rogalski, E. (2013). A novel frontal pathway underlies verbal fluency in primary progressive aphasia. *Brain*, 136(8), 2619–2628. <https://doi.org/10.1093/brain/awt163>
- Costa, A., Bagoj, E., Monaco, M., Zabberoni, S., De Rosa, S., Papanonio, A. M., Mundi, C., Caltagirone, C., & Carlesimo, G. A. (2014). Standardization and normative data obtained in the Italian population for a new verbal fluency instrument, the phonemic/semantic alternate fluency test. *Neurological Sciences*, 35(3), 365–372. <https://doi.org/10.1007/s10072-013-1520-8>
- Crowe, S. F. (1998). Decrease in performance on the verbal fluency test as a function of time: Evaluation in a young healthy sample. *Journal of clinical and experimental neuropsychology*, 20(3), 391–401. <https://doi.org/10.1076/jcen.20.3.391.810>
- Demetriou, E., & Holtzer, R. (2017). Mild cognitive impairments moderate the effect of time on verbal fluency performance. *Journal of the International Neuropsychological Society*, 23(1), 44–55. <https://doi.org/10.1017/S1355617716000825>
- Diamond, A. (2013). Executive functions. *Annual review of psychology*, 64, 135–68. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Dursun, S. M., Robertson, H. A., Bird, D., Kutcher, D., & Kutcher, S. P. (2002). Effects of ageing on prefrontal temporal cortical network function in healthy volunteers as assessed by COWA. *An exploratory survey. Progress in Neuropsychopharmacology Biological Psychiatry*, 26(5), 1007–1010. [https://doi.org/10.1016/s0278-5846\(01\)00321-9](https://doi.org/10.1016/s0278-5846(01)00321-9)
- Faroqi-Shah, Y., & Milman, L. (2018). Comparison of animal, action and phonemic fluency in aphasia. *International Journal of Language & Communication Disorders*, 53(2), 370–384. <https://doi.org/10.1111/1460-6984.12354>
- Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language*, 115(3), 149–161. <https://doi.org/10.1016/j.bandl.2010.07.006>
- Fortuny, L. A., Heaton, R. K., & Hermosillo, D. (1998). Neuropsychological comparisons of Spanish-speaking participants from the US-Mexico border region versus Spain. *Journal of the International Neuropsychological Society*, 4(4), 363–379. <https://doi.org/10.1017/S1355617798003634>
- Henry, J. D., & Phillips, L. H. (2006). Covariates of production and perseveration on tests of phonemic, semantic and alternating fluency in normal aging. *Aging, Neuropsychology, and Cognition*, 13(3-4), 529–551. <https://doi.org/10.1080/138255890969537>
- Herbert, V., Brookes, R. L., Markus, H. S., & Morris, R. G. (2014). Verbal fluency in cerebral small vessel disease and Alzheimer's disease. *Journal of the*

- International Neuropsychological Society*, 20(4), 413–421. <https://doi.org/10.1017/S1355617714000101>
- Ivnik, R. J., Malec, J. F., Smith, G. E., Tangalos, E. G., & Petersen, R. C. (1996). Neuropsychological tests' norms above age 55: COWAT, BNT, MAE token, WRAT-R reading, AMNART, STROOP, TMT, and JLO. *The Clinical Neuropsychologist*, 10, 3, 262–278. <https://doi.org/10.1080/13854049608406689>
- Jacobs, S., Mercuri, G., & Holtzer, R. (2021). Assessing within-task verbal fluency performance: The utility of individual time intervals in predicting incident mild cognitive impairment. *Aging, Neuropsychology, and Cognition*, 28(5), 733–747. <https://doi.org/10.1080/13825585.2020.1817306>
- Johnson-Selfridge, M., & Zalewski, C. (2001). Moderator variables of executive functioning in schizophrenia: Meta-analytic findings. *Schizophrenia Bulletin*, 27(2), 305–316.
- Khalil, M. S. (2010). Preliminary Arabic normative data of neuropsychological tests: The verbal and design fluency. *Journal of Clinical and Experimental Neuropsychology*, 32(9), 1028–1035. <https://doi.org/10.1080/13803391003672305>
- Kozora, E., & Cullum, C. M. (1995). Generative naming in normal aging: Total output and qualitative changes using phonemic and semantic constraints. *The Clinical Neuropsychologist*, 9(4), 313–320. <https://doi.org/10.1080/13854049508400495>
- Kraan, C., Stolwyk, R. J., & Testa, R. (2013). The abilities associated with verbal fluency performance in a young, healthy population are multifactorial and differ across fluency variants. *Applied Neuropsychology: Adult*, 20(3), 159–168. <https://doi.org/10.1080/09084282.2012.670157>
- Lee, P. H. (2014). Is a cutoff of 10% appropriate for the change-in-estimate criterion of confounder identification? *Journal of Epidemiology*, 24(2), 161–167. <https://doi.org/10.2188/jea.JE20130062>
- Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th edn). Oxford University Press.
- Lobo, A., Saz, P., Marcos, G., Día, J. L., de la Cámara, C., Ventura, T., Asín, F., Pascual, L. F., Montañés, J. A., & Aznar, S. (1999). Revalidación y normalización del mini-Examen Cognoscitivo (primera versión en castellano del mini-mental status examination) en la población general geriátrica [re-validation of the mini-Examen Cognoscitivo (first Spanish version of the mini-mental status examination) and population-based norms in the elderly community]. *Medicina Clínica*, 112(20), 767–774.
- Loonstra, A. S., Tarlow, A. R., & Sellers, A. H. (2001). COWAT metanorms across age, education, and gender. *Applied Neuropsychology*, 8(3), 161–166. [https://doi.org/10.1207/S15324826AN0803\\_5](https://doi.org/10.1207/S15324826AN0803_5)
- López-Higes, R., Rubio-Valdehita, S., Fernández-Blázquez, M. A., Lojo-Seoane, C., Ávila-Villanueva, M., Montenegro-Peña, M., Mallo, S. C., & Delgado-Losada, M. L. (2022). Spanish Consortium for Ageing Normative Data (SCAND): semantic verbal fluency tests. *Archives of Clinical Neuropsychology*, 37(2), 352–364. <https://doi.org/10.1093/arclin/acab059>
- Lubrin, G., Periañez, J. A., Laseca-Zaballa, G., Bernabéu-Brotos, E., & Ríos-Lago, M. (2022). Verbal fluency tasks: influence of age, gender, and education and normative data for the Spanish native adult population. *Archives of Clinical Neuropsychology*, 37(2), 365–375. <https://doi.org/10.1093/arclin/acab056>
- 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 Boston naming test, controlled oral word association, category fluency, animal naming, token test, wrat-3 reading, trail making test, Stroop test, and judgment of line orientation. *The Clinical Neuropsychologist*, 19(2), 243–269. <https://doi.org/10.1080/13854040590945337>
- Marquine, M. J., Morlett Paredes, A., Madriaga, C., Blumstein, Y., Umlauf, A., Kamalyan, L., & Cherner, M. (2021). Demographically-adjusted norms for selected tests of verbal fluency: Results from the Neuropsychological Norms for the US-Mexico Border Region in Spanish (NP-NUMBRS) project. *The Clinical Neuropsychologist*, 35(2), 269–292. <https://doi.org/10.1080/13854046.2020.1762931>
- Mathuranath, P. S., George, A., Cherian, P. J., Alexander, A., Sarma, S. G., & Sarma, P. S. (2003). Effects of age, education and gender on verbal fluency. *Journal of Clinical and Experimental Neuropsychology*, 25(8), 1057–1064. <https://doi.org/10.1076/jcen.25.8.1057.16736>
- McDowd, J., Hoffman, L., Rozek, E., Lyons, K. E., Pahwa, R., Burns, J., & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychology*, 25(2), 210–25. <https://doi.org/10.1037/a0021531>
- Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24(1), 167–202. <https://doi.org/10.1146/annurev.neuro.24.1.167>
- Mitrushina, M., Boone, K. B., Razani, J., & D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment*. New York, NY: Oxford University Press.
- Moraes, A. L., Guimarães, L. S. P., Joannette, Y., Parente, M. A. D. M. P., Fonseca, R. P., & de Almeida, R. M. M. (2013). Effect of aging, education, reading and writing, semantic processing and depression symptoms on verbal fluency. *Psicologia Reflexao e Critica*, 26(4), 680–690. <https://doi.org/10.1590/S0102-79722013000400008>
- Noble, K. G., Houston, S. M., Brito, N. H., Bartsch, H., Kan, E., Kuperman, J. M., & Akshoomoff, N. Amaral, D. G. Bloss, C. S. Libiger, O. Schork, N. J. Murray, S. S. Casey, B. J. Chang, L. Ernst, T. M. Frazier, J. A. Gruen, J. R. Kennedy, D. N. Van Zijl, P. . . . Sowell, E. R. (2015). Family income, parental education and brain structure in children and adolescents. *Nature Neuroscience*, 18(5), 773–778. <https://doi.org/10.1038/nn.3983>
- Oberg, G., & Ramírez, M. (2006). Cross-linguistic meta-analysis of phonological fluency: Normal performance across cultures. *International Journal of Psychology*, 41(5), 342–347. <https://doi.org/10.1080/00207590500345872>
- Pakhomov, S., Eberly, L. E., & Knopman, D. S. (2018). Recurrent perseverations on semantic verbal fluency tasks as an early marker of cognitive impairment. *Journal of Clinical and Experimental Neuropsychology*, 40(8), 832–840. <https://doi.org/10.1080/13803395.2018.1438372>
- Panico, F., Sagliano, L., Magliacano, A., Santangelo, G., & Trojano, L. (2022). The relationship between cognitive reserve and cognition in healthy adults: a systematic review. *Current Psychology*, 1–13. <https://doi.org/10.1007/s12144-022-03523-y>
- Pekkala, S. (2012). Verbal fluency tasks and the neuropsychology of language. In Faust, M. (Eds.), *The Handbook of the Neuropsychology of Language*. 619, 634, Wiley-Blackwell.
- Peña-Casanova, J., Quiñones-Ubeda, S., Gramunt-Fombuena, N., Quintana-Aparicio, M., Aguilar, M., & Badenes, D. Spanish Multicenter Normative Studies (NEURONORMA Project): Norms for verbal fluency tests. *Archives of Clinical Neuropsychology*, 24(4), 395–411. <https://doi.org/10.1093/arclin/acp042>
- Pereira, A. H., Goncalves, A. B., Holz, M., Goncalves, H. A., Kochhann, R., Joannette, Y., Zimmermann, N., & Fonseca, R. P. (2018). Influence of age and education on the processing of clustering and switching in verbal fluency tasks. *Dementia e Neuropsychologia*, 12(4), 360–367. <https://doi.org/10.1590/1980-57642>
- Pérez, J., Abellán, A., Aceituno, P., & Ramiro, D. (2020). Un perfil de las personas mayores en España, Indicadores estadísticos básicos. *Informes Envejecimiento en Red*, 25, 1–39, 2020, <http://envejecimiento.csic.es/documentos/documentos/enred-indicadoresbasicos2020.pdf>
- Pérez, J., Ramiro, D., Aceituno, P., Muñoz Diaz, C., Bueno, C., Ruiz-Santacruz, J., . . . & Villuendas, B. (2022). Un perfil de las personas mayores en España, Indicadores estadísticos básicos. *Madrid, Informes Envejecimiento en red* no, 29, 40.
- Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, 256(3), 183–194.
- Raboutet, C., Sauzéon, H., Corsini, M. M., Rodrigues, J., Langevin, S., & N'kaoua, B. (2010). Performance on a semantic verbal fluency task across time: Dissociation between clustering, switching, and categorical exploitation processes. *Journal of Clinical and Experimental Neuropsychology*, 32(3), 268–280. <https://doi.org/10.1080/13803390902984464>
- Rivera Mindt, M., Marquine, M. J., Aghvinian, M., Paredes, A. M., Kamalyan, L., Suárez, P., & Cherner, M. (2021). The neuropsychological norms for the US-Mexico border region in Spanish (NP-NUMBRS) project: Overview and considerations for life span research and evidence-based practice. *The Clinical Neuropsychologist*, 35(2), 466–480. <https://doi.org/10.1080/13854046.2020.1794046>
- Rofes, A., De Aguiar, V., Ficek, B., Wendt, H., Webster, K., & Tsapkini, K. (2019). The role of word properties in performance on fluency tasks in people with primary progressive aphasia. *Journal of Alzheimer's Disease*, 68(4), 1521–1534. <https://doi.org/10.3233/JAD-180990>

- Roldán-Tapia, L., García, J., Cánovas, R., & León, I. (2012). Cognitive reserve, age, and their relation to attentional and executive functions. *Applied Neuropsychology: Adult*, 19(1), 2–8. <https://doi.org/10.1080/09084282.2011.595458>
- Santos Nogueira, D., Azevedo Reis, E., & Vieira, A. (2016). Verbal fluency tasks: effects of age, gender, and education. *Folia Phoniatica et logopaedica: Official Organ of the International Association of Logopedics and Phoniatrics (IALP)*, 68(3), 124–133. <https://doi.org/10.1159/000450640>
- Sauzéon, H., Raboutet, C., Rodrigues, J., Langevin, S., Schelstraete, M. A., Feyereisen, P., & N'Kaoua, B. (2011). Verbal knowledge as a compensation determinant of adult age differences in verbal fluency tasks over time. *Journal of Adult Development*, 18(3), 144–154. <https://doi.org/10.1007/s10804-010-9107-6>
- Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, 5, 772. <https://doi.org/10.3389/fpsyg.2014.00772>
- Shores, E., Carstairs, J., & Crawford, J. (2006). Excluded Letter Fluency Test (ELF): Norms and test-retest reliability data for healthy young adults. *Brain Impairment*, 7(1), 26–32. <https://doi.org/10.1375/brim.7.1.26>
- Smith, E. E., Smith, J. A., & Juengst, S. B. (2020). Cognitive process scores associated with self-reported behavioral dysfunction on the Frontal Systems Behavior Scale (FrSBe) in chronic traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, 42(1), 90–100. <https://doi.org/10.1080/13803395.2019.1676882>
- Strauss, E., Sherman, E. M. S., Spreen, O., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed). Oxford University Press.
- Thiele, K., Quinting, J. M., & Stenneken, P. (2016). New ways to analyze Word generation performance in brain injury: A systematic review and meta-analysis of additional performance measures. *Journal of Clinical and Experimental Neuropsychology*, 38(7), 764–781. <https://doi.org/10.1080/13803395.2016.1163327>
- Tombaugh, T. N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology*, 14(2), 167–177. [https://doi.org/10.1016/S0887-6177\(97\)00095-4](https://doi.org/10.1016/S0887-6177(97)00095-4)
- Venegas, M. J., & Mansur, L. L. (2011). Verbal fluency: Effect of time on item generation. *Dementia & Neuropsychologia*, 5(2), 104–107. <https://doi.org/10.1590/S1980-57642011DN05020008>
- Yesavage, J., Brink, T., Rose, T., Lum, Q., Huang, Y., Adey, M., & Leirer, V. O. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17, 37–49. [https://doi.org/10.1016/0022-3956\(82\)90033-4](https://doi.org/10.1016/0022-3956(82)90033-4)