Language and Cognition (2025), 17, e76, 1–22
doi:10.1017/langcog.2025.10033

ARTICLE

Verbal fluency with high executive load for early dementia's risk detection

Ramón López-Higes[1], Sara M. Fernandes[2], Pedro F. S. Rodrigues[3], Barbara Paiva[4], Susana Rubio-Valdehita[5] and David Del Río[1]

[1]Department of Experimental Psychology, Complutense University of Madrid, Madrid, Spain; [2]CINTESIS.UPT@RISE-HEALTH, Portucalense University, Porto, Portugal; [3]CINTESIS.UPT@RISE-HEALTH, Portucalense University, Porto, Portugal; [4]CINTESIS.UPT@RISE-HEALTH, Portucalense University, Porto, Portugal and [5]Department of Social, Work and Differential Psychology, Complutense University of Madrid, Madrid, Spain
Corresponding author: Susana Rubio-Valdehita; Email: srubiova@ucm.es

(Received 12 March 2025; Revised 23 August 2025; Accepted 07 September 2025)

Abstract
Verbal fluency (VF) tasks are used in cognitive assessments to detect early signs of neurodegenerative diseases like Alzheimer's. This study aimed to assess the contribution of VF tasks with varying executive processing loads to the early identification of cognitive impairment in the preclinical stage of subjective cognitive decline (SCD). A total of 97 older adults were classified into three groups: healthy controls (HC), SCD and mild cognitive impairment (MCI). Participants completed phonemic, semantic, alternating and orthographic VF tasks. Education level significantly affected VF performance, with gender differences being inconsistent. The HC and SCD groups performed similarly in phonemic and semantic tasks but differed significantly in high-executive-load tasks, where SCD participants performed worse. MCI patients showed lower performance across all VF tasks. Discriminant and ROC analyses identified alternating and orthographic VF tasks as effective markers for distinguishing cognitive status, supporting their potential for early detection of Alzheimer's disease.

Keywords: executive load; healthy older adults; individual variables; mild cognitive impairment; neuropsychological assessment; subjective cognitive decline; verbal fluency

1. Introduction

Verbal fluency (VF) tasks are fundamental instruments in the assessment of cognitive functioning (Demetriou & Holtzer, 2017), specifically in the elderly, due to their ability to identify the speed of access to the mental lexicon, to assess its integrity, as well as the integrity of semantic memory. VF tasks also involve executive functions,

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

such as updating, inhibitory control, cognitive flexibility and self-regulation (Aiello et al., 2022; Costa et al., 2014; Henry, Crawford & Phillips, 2004; Lam & Marquardt, 2022; Shao, Janse, Visser & Meyer, 2014; Troyer, 2000). There are four VF variants (grammatical, phonological, categorical/semantic and combined forms), a specific amount of time to do the task (usually, 60 seconds) and rules defining what cannot be done during evocation.

Grammatical VF is related to the evocation of adjectives or verbs during a limited time (Piatt, Fields, Paolo & Tröster, 2004). The most used forms of VF tasks are the phonological VF (PVF in advance) and the semantic variant (SVF). PVF requires the production of words that begin with a specific letter (e.g., A), whereas SFV requires the production of words belonging to a specific semantic category (e.g., animals) (Demetriou & Holtzer, 2017). PVF is thought to primarily involve controlled access to memory representations in the mental lexicon, with support from the frontal lobe (Baldo, Schwartz, Wilkins & Dronkers, 2010; Marko et al., 2023). The search process in SVF mirrors the natural way of searching and retrieving information from the lexicon (Shao, Janse, Visser & Meyer, 2014) and consequently reflects automatic associative retrieval processes related to the integrity of the temporal lobes. Thus, the search process and performance in SVF, in addition to executive functions, rely on the organization of semantic memory (Henry, Crawford & Phillips, 2004). In contrast, PVF must adhere to artificial constraints, which require the suppression of our natural retrieval methods (Heim, Eickhoff, Friederici & Amunts, 2009). This is why PVF is considered to have greater executive load; thus, many researchers view these tasks as valuable tools that are particularly sensitive in identifying cognitive disorders affecting the frontal lobe (Cipolotti et al., 2018; Rami, et al, 2007).

The cognitive mechanisms underlying each task are difficult to delineate because VF is multidimensional and maps onto cognitive domains including language, executive functions and processing speed. For example, Kiselica, Webber, and Benge (2020) reported that PVF loaded on executive functioning/speed factors, while SVF loaded on language factors. In other studies, both PVF and SVF are loaded on the language factor with limited support for an executive contribution to either type (Dowling, Hermann, La Rue & Sager, 2010; Whiteside et al., 2016). More recently, Aita et al. (2019) found that both tasks loaded onto a unique executive factor.

A special mention must be made for two combined VF forms: alternating SVF and PVF tasks (Costa et al, 2014; Lam & Marquardt, 2022), and SVF with orthographic constraint (Shores, Carstairs & Crawford, 2006). Alternating VF, which requires shifting between semantic categories or letters throughout a task, involves additional attentional resources to track and prepare for upcoming semantic categories or phonemic targets. But, more importantly, shifting during alternating VF requires active suppression of clustering strategies, which should lead to greater cognitive effort compared to non-alternating VF (Lam & Marquardt, 2022). Compared to typical fluency tasks, the orthographic constraint VF task is assumed to require more inhibitory control abilities (Macoir, Tremblay & Hudon, 2022). To our knowledge, Macoir et al. (2022) is the only previous study showing that VF tasks with higher executive load are useful for detecting cognitive deficits in the preclinical stage of AD. However, Macoir et al. did not compare alternating VF and orthographic constraint VF to assess their level of difficulty (executive processing load). In the absence of prior empirical data establishing their relative difficulty, we assume that these FV variants are functionally comparable in terms of their executive demands.

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

Given all these arguments, it seems possible to establish a kind of gradient regarding the executive processing loads of FV tasks: SVF < PVF < Alternating FV = Orthographic Restriction FV (the equal sign indicates that they are functionally comparable in terms of executive demands).

Several factors, including age, educational level and gender, significantly influence performance on VF tasks, with education being the most impactful (Ekström, Josefsson, Bäckman & Laukka, 2024; Hirnstein, Stuebs, Moè & Hausmann, 2023; Delgado-Losada et al., 2024; López-Higes et al., 2022; Lubrini et al., 2022; Peña-Casanova et al., 2009). Most studies find no gender differences in PVF, although women tend to perform better than men in SVF for certain categories (Costa et al., 2014; López-Higes et al., 2022; Peña-Casanova et al., 2009). A decline in performance with age is reported in many studies (Cavaco et al., 2013; Contador et al., 2016; López-Higes et al., 2022; Peña-Casanova et al., 2009), although some recent research did not find age effects on PVF tasks (Delgado-Losada et al., 2024). In the Portuguese population, age and education consistently affect both PVF and SVF. Gender effects are inconsistent, but when observed, it tend to favor women in certain categories (Cavaco et al., 2013; Santos Nogueira et al., 2017; Vicente et al., 2022).

The performance on a VF test is usually evaluated by using the total number of correct words given within the time limit (Pekkala, 2012). However, different authors have proposed more informative measures. Troyer et al. (1997) proposed for the SVF tasks the mean cluster size and the number of switches between clusters. The average cluster size is associated with the word production process within a given semantic group/field. Basically, each cluster consists of two or more consecutive words belonging to a specific subcategory (e.g., domestic animals). The number of switches reflects, as its name suggests, the process of shifting from one semantic subcategory to another. It is obtained by counting the number of times each participant switches from one subcategory to another during the task.

In some studies, time segments have also been considered, provided that, in general, participants produce most of the words in the early stages of the task (first 20–30 seconds) using a semi-automatic rapid retrieval process (Linz et al., 2019; Venegas & Mansur, 2011). Recent studies have computed the so-called 'semantic-phonemic discrepancy' (Aiello et al., 2023) by subtracting PVF scores from SVF ones (positive and negative values corresponding to a semantic and phonemic advantage, respectively), which seems to predict tauopathy across the AD spectrum.

The important role played by executive functions in VF has been highlighted in the preceding lines. We also know that impaired executive function is considered a sign of cognitive decline in the prodromal phase of AD (Bairami et al., 2023; García-González, 2024; Nevado et al., 2021). Early identification of neurodegenerative processes is a critical concern for public health as well as clinical and scientific research, since preventing the appearance of signs and symptoms of dementia and the diagnosis of AD is essential to reduce its negative impacts (McDonough & Allen, 2019).

The typical progression of neurocognitive disorders associated with Alzheimer's disease (AD) can be understood as a continuum from normal aging to dementia. It begins with an asymptomatic phase, where individuals show no noticeable symptoms, progressing into subjective cognitive decline (SCD). In this early stage, individuals experience subjective complaints about their memory or cognition, which are linked to an increased risk of future cognitive impairment and AD development (Jessen et al., 2014). Although diagnosing SCD is challenging due to its subjective

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

nature, criteria have been established to identify it, emphasizing the importance of early detection. Evidence suggests that some individuals with SCD may be in a preclinical stage of AD, especially when associated with biomarkers like amyloid deposits (Sanabria et al., 2018) or subtle cognitive deficits such as reduced semantic VF (Nikolai et al., 2018). As the condition advances, some individuals develop mild cognitive impairment (MCI), particularly the amnestic subtype, characterized by objective memory deficits but preserved daily functioning. MCI represents an intermediate state between normal aging and dementia, with a higher likelihood of progression to AD. Diagnostic criteria for MCI include subjective complaints confirmed by informants, measurable cognitive decline and maintained independence in daily activities, although subtle impairments may be present (Demetriou & Holtzer, 2017; Petersen, 2004; Petersen et al., 2014). Finally, when cognitive decline becomes substantial, interfering with independence and involving multiple domains, a diagnosis of Alzheimer's dementia is made. This stage is marked by significant impairments in memory, reasoning, language and other cognitive functions, often confirmed through neuropsychological testing and supported by biomarkers such as neuroimaging and cerebrospinal fluid (CSF) analysis (Jack et al., 2018).

In this context, an important challenge could be to determine which VF tasks can serve as early markers of a neurodegenerative process in the preclinical stage of the disease. In this sense, some longitudinal studies have revealed that patients with preclinical AD show a differentially worse decline in SVF compared to PVF, highlighting the sensitivity of SVF for detecting AD progression in the early stages (Demetriou & Holtzer, 2017; Wright, De Marco & Venneri, 2023). Other cross-sectional studies have reported that individuals with SCD, who are in an intermediate preclinical stage in the continuum from normality to dementia, show lower performance on SVF tasks compared to cognitively intact individuals (Koppara et al., 2015; Nikolai et al., 2018). A recent study conducted by Macoir et al. (2022) concluded that high-executive-processing-load VF tasks are useful for the identification of cognitive deficits in SCD.

To determine the potential contribution of VF to the early identification of cognitive impairment in the preclinical stage of AD, the present study explored the performance of older adults with SCD, MCI and healthy controls (HCs) in VF tasks having different executive processing loads. The HC group will show performance across the VF tasks according to the previously mentioned gradient: SVF < PVF < Alternating VF = Arthographic Constraint VF. By hypothesis, the group of participants with SCD is expected to have a performance similar to that of the HC group in both the SVF and PVF tasks, and a significantly lower performance in the tasks with high executive load. Finally, the performance of the MCI patient group is expected to be significantly lower than that of the HC group in all VF tasks, and in addition, in the case of high-executive-load tasks, it will also be lower than that of the SCD group.

2. Methodology

2.1. Participants

The total sample consisted of 97 participants from the northern region of Portugal, with an average age of 77.40 years ($SD = 5.54$), distributed by gender (47 females and 50 males). The recruitment process was conducted through direct contact with the

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

heads or directors of senior universities and day social centers, who were informed about the study's objectives and characteristics. This aimed to reach participants with different cognitive profiles and sociodemographic characteristics, such as educational level. After the initial contact, a letter and/or an email were delivered to the heads or directors of the centers who were interested. These included basic information about the study, the call for volunteer participants in the research and different forms of contact with project personnel.

The educational level of the participants is varied, although primary education predominates as the main category: 1% only have basic literacy (basic ability to read and write), 67% primary education, 21.6% have completed the second and third cycles, 7.2% have secondary education and 3.1% reached higher education. Participants were divided into three groups: HC elders, SCD and patients with MCI according to the following exclusion and inclusion criteria.

Subjects with a prior diagnosis of neurological or psychiatric diseases or having severe symptoms at the starting point of the study were excluded. Furthermore, those who had problems with alcohol or were under the influence of medications that could influence cognitive performance, as well as those who had levels of depressive or anxious symptoms above the normative cutoff point: 4 on the Geriatric Depression Scale (GDS-15) (Apostle et al., 2014) for depression, and 32 on the State–Trait Anxiety Inventory (STAI-Y2) (Silva & Spielberger, 2007) for anxiety, were also excluded from the study.

According to Jessen et al.'s (2014) criteria, SCD participants have a self-perception of progressive deterioration in cognitive functioning (mainly associated with memory) and have memory or other subjective cognitive complaints. To determine the presence of memory complaints, the following three questions were used:

1. Do you have memory or other cognitive problems?
   Yes/No.
2. Have you lost memory in recent years?
   a) No.
   b) Yes, but not much, it is not a problem because it does not affect my daily life.
   c) Yes, and that is a problem for me because it somewhat affects my daily life.
   d) Yes, and it causes me significant problems because it significantly affects my daily life.
3. Have you consulted a doctor or any professional about your memory problems?
   Yes, who have you consulted?
   No.

The participants were initially categorized as having SCD if they responded yes to question 1, selected alternatives c or d in question 2 and answered yes in question 3 (sometimes people answer no to this question). Participants' responses to the Everyday Memory Questionnaire (EMQ) (Ribeiro et al., 2024) evidenced memory failures (greater than 10 points). All of them had normal scores on the Auditory Verbal Learning Test (AVLT) (Cavaco et al., 2015) (difference between fifth and first essays [learning]; delayed recall), in the Memory Strategies Test scores (TEM) (Fernandes et al., 2018) and preserved activities of daily living according to the Functional Scale of Instrumental Activities of Daily Living (IADL) (Lawton, 1969; Núcleo de Estudos de Geriatria da Sociedade Portuguesa de Medicina Interna – GERMI, 2017). Montreal Cognitive Assessment Test (MOCA) scores denote

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

normality according to Freitas et al. norms adjusted for age and educational level (Freitas et al., 2010). The group with SCD consisted of 33 participants (14 women and 16 men), with a mean age of 76.81 (*SD* = 5.60). In this group, 50% of the participants have basic/primary education, 31.3% have completed the second and third cycles, 15.6% have secondary education and 3.1% have higher education.

The inclusion criteria for participants with MCI (Petersen, 2004; 2011) were based on corroborated memory complaints using the three questions described above, and the above-mentioned cut-off point in EMQ, evidencing significant memory failures. Additionally, they showed a deficit in the memory domain assessed by AVLT and TEM. All participants in this group have sufficiently preserved ability to perform activities of daily living, as evidenced by the IADL, and they do not meet diagnostic criteria for dementia. MOCA scores of these participants are below the mean, considering age and educational level (Freitas et al., 2010). The group of patients with MCI had a total of 34 participants (16 women and 17 men), with a mean age of 79.65 (*SD* = 5.43). In terms of education, 2.9% have only basic literacy, 94.1% have completed primary education and 2.9% have completed the second and third cycles.

Identification of the HC group involved the absence of complaints related to memory, language or other cognitive functions, in addition to presenting normal or even above-average performance in memory tests (AVLT and TEM) and IADL. MOCA scores denote normality according to Freitas et al. (2010) norms adjusted for age and educational level. The HC group had a total of 31 participants (16 women and 15 men), with a mean age of 75.55 (*SD* = 4.88). This group had 54.8% participants with primary education, 32.3% had completed the second and third cycles, 6.5% with secondary education and 6.5% with higher education.

2.2. Instruments

Initially, a sociodemographic questionnaire was applied to collect basic information about participants, such as age, gender, education level, medical history and other relevant variables. A set of instruments was selected to classify participants according to the inclusion/exclusion criteria mentioned before:

a) EMQ (Rodrigues et al., 2025; Royle & Lincoln, 2008) is an instrument that assesses memory failure in everyday life. It contributes to a deeper understanding of the memory difficulties perceived by the subjects. Evidence of a strong internal reliability (Cronbach's alpha of 0.92) of the scale items (Rodrigues et al., 2025). The use of this instrument made it possible to differentiate participants who presented subjective memory complaints.
b) AVLT (Cavaco et al., 2015; Ryan & Geisser, 1986) is a relevant instrument for evaluating verbal memory in elderly people, especially those with cognitive dysfunctions. This instrument integrates immediate and delayed recall tests and provides information on verbal acquisition, learning and retention (Cavaco et al., 2015). It has excellent internal consistency (Cronbach's alpha of 0.89). This instrument is crucial to highlight deficits in the memory domain and to differentiate the performance of participants with MCI, SCD and HC.
c) TEM (Fernandes et al., 2018; Yubero, Maestú, Paul & Gil, 2011) is an instrument that assesses the impact of executive functions in declarative memory tasks, distinguishing whether memory deficits are due to primary memory problems or executive difficulties. Useful for evaluating memory and

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

executive functions in different age groups, presenting criterion validity and providing empirical data for interpreting performance in memory tests. TEM has acceptable internal consistency (Cronbach's alpha of 0.74) (Fernandes et al, 2018).

d) IADL (Lawton, 1969; Núcleo de Estudos de Geriatria da Sociedade Portuguesa de Medicina Interna – GERMI, 2017) is an instrument that assesses the necessary skills so that the elderly can live in a healthy way independently. It is essential for identifying changes in functional capacity over time, providing insights into an individual's daily functioning across eight domains of function (Graf, 2008). This instrument shows high internal consistency (Cronbach's alpha of 0.88) in its item scale (Araújo, Paúl & Ribeiro, 2017). This instrument was relevant in this study to assess the level of independence of the elderly and thus exclude individuals who do not have sufficiently preserved autonomy for daily activities.
e) STAI-Y2 (Silva & Spielberger, 2007; Spielberger et al., 1971) is a self-report questionnaire to assess state–trait anxiety using a Likert scale. It is extremely used and reliable across different age groups, with good internal consistency (Cronbach's alpha of 0.87) (Bieling, Antony & Swinson, 1998; Julian, 2011). This instrument identifies the presence or absence of anxious symptoms, thus allowing the exclusion of participants who present levels of anxious symptoms above the cutoff point.
f) GDS (Apostle, et al., 2014; Yesavage, et al., 1982) is an instrument that identifies the presence and severity of depressive symptoms. It consists of a series of questions that address feelings and behaviors associated with depression, revealing good internal consistency (Cronbach's alpha of 0.83) (Apostle, et al., 2014; Yesavage, et al., 1982). GDS is essential to control depressive symptoms that can interfere with participants' cognition, so it is used to exclude participants who have a level of depression above the cutoff point.
g) MOCA (Freitas et al., 2010; Nasreddine et al., 2005) is an effective and practical cognitive screening instrument to detect cognitive impairment. This instrument assesses six cognitive domains: memory, executive functions, attention and concentration, language, visuospatial ability and orientation. Furthermore, it shows good internal consistency in the Portuguese adaptation, with a Cronbach's alpha coefficient of 0.79 (Freitas et al., 2010).

The study includes in its experimental phase four types of VF tasks with different executive loads. In the PVF tasks, participants evoke words starting with the letter 'A', and then with the letter 'S'. In SVF tasks, participants must produce animal names and then pieces of clothing. In the alternating VF variant, the subjects were asked to alternate between words starting with the letter 'M' and then a name of a fruit category, continuing in the same manner until the end of the established time. In the orthographic constraint VF, participants must produce names corresponding to parts of the body that did not have the letter 'R' until the end of time.

2.3. Procedure

The study procedure began with sending the project proposal to the Health Ethics Committee of the Portucalense University Infante D. Henrique. The data collection process began with the prior sending of authorization requests to the participant

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

institutions to carry out the collection on their premises, requesting the guarantee of an adequate and quiet environment, with a table and two chairs. After acceptance by the institutions, informed consent was presented to the voluntary participants with details on the purpose of the study, the procedures involved, the exclusion criteria, the rights of the participants and the guarantee of confidentiality of the information provided. Participants had the opportunity to review informed consent and ask questions before agreeing to participate. Data collection began only after the participant signed the informed consent. Furthermore, the data collection process was carried out ethically, guaranteeing the privacy and confidentiality of the participants.

This process was individual and carried out in a single moment, divided into two blocks separated by a break, with an average duration of 60–90 minutes, depending on the participant. The first block consisted of the application of the neuropsychological assessment protocol described in the Instruments subsection, adapted to the Portuguese population.

After a break, all participants were subjected to an experimental phase in which they performed the four VF tasks. All these tasks had an execution time of 90 seconds each. The order of presentation of these tasks was counterbalanced across subjects.

When participants do not follow the instructions, lose the thread or forget what to do, the examiner provides gentle reminders or prompts to help them refocus on the task. If they continue to have difficulty, the examiner might give a brief rest period or clarify the instructions again. In some cases, if the participant is unable to continue despite prompts, the task is discontinued, and the performance is recorded as incomplete.

2.4. Data analysis

As a previous step, we computed mean scores for PVF and SVF tasks. Descriptive statistics were included for sociodemographic variables and neuropsychological characteristics of groups. Then, a multivariate ANOVA was used to find significant differences between the groups in all these variables.

In the second phase, we obtained descriptive statistics for the groups across tasks and time intervals, as well as the correlations between participants' years of education and all dependent measures. A generalized mixed ANOVA (Group [3] x Task [4] x Time interval [3]), with participants' years of education as a covariate, was conducted to search for significant simple and interaction effects. Partial Eta square (η^2_p) was used to estimate the effect size. A series of ANOVAs was used to explore the significant interactions involving Group. Paired-sampled *t* tests were also used to search for differences between the pairs (1) PVF and SVF, and (2) alternating and orthographic constraint VF in the HC group.

A discriminant analysis was performed to determine the relative weight of VF variants in subjects' classification. Finally, ROC curve analyses were performed to check if the variables that had the most important role in classification should show individually good diagnostic accuracy and to obtain cut-off values with potential clinical significance in those variables.

3. Results

3.1. Descriptives and differences between groups

A multivariate ANOVA with all the variables summarized in Table 1 showed that there were statistically significant differences between the three groups in the

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

Table 1. Descriptive data by group and statistical differences between groups

	Group	Mean	Std. Deviation	HC	SCD	MCI
Age	HC	75.55	4.88			
	SCD	76.58	5.54			
	MCI	78.37	4.29			
Years of education	HC	6.58	3.29			*
	SCD	7.16	3.32			**
	MCI	4.03	1.03	*	**	
EMQ	HC	1.84	1.04		**	**
	SCD	12.42	5.11	**		
	MCI	11.60	6.78	**		
Learning_AVLT	HC	4.87	2.01			**
	SCD	4.65	2.73			**
	MCI	1.97	.93	**	**	
AVLT_delayed recall	HC	6.74	3.28			**
	SCD	5.81	2.18			**
	MCI	1.20	.96	**	**	
TEM1	HC	3.61	1.15			**
	SCD	3.45	.99			**
	MCI	1.40	.72	**	**	
TEM2	HC	4.13	1.06			**
	SCD	3.84	1.70			**
	MCI	2.13	.57	**	**	
TEM3	HC	5.16	.97			**
	SCD	5.58	1.67			**
	MCI	2.87	.68	**	**	
TEM4	HC	6.61	1.33			**
	SCD	6.03	1.47			**
	MCI	3.40	.62	**	**	
TEM5	HC	6.23	1.43			**
	SCD	6.58	1.50			**
	MCI	4.00	.59	**	**	
IADL	HC	6.58	1.48			
	SCD	6.32	1.42			
	MCI	5.90	1.32			
STAI_Y2	HC	31.84	2.31			
	SCD	30.77	2.06			
	MCI	31.57	1.81			
GDS_15	HC	2.71	1.04			
	SCD	2.94	.89			
	MCI	2.83	.95			
MOCA	HC	24.00	2.13			**
	SCD	23.10	1.76			**
	MCI	18.60	.50	**	**	

Notes: * p < .01; ** p < .001.

Abbreviations: EMQ: Everyday Memory Questionnaire; AVLT: Auditory Verbal Learning Test; TEM: Test de Estrategias de Memoria (English trad.: Memory Strategies Test); IADL: Functional Scale of Instrumental Activities of Daily Living; STAI-Y2: State–Trait Anxiety Inventory; GDS-15: Geriatric Depression Scale 15 items; MOCA: Montreal Cognitive Assessment Test.

following variables: Years of education ($F[2, 89] = 10.88$, $p < .001$), EMQ ($F[2, 89] = 44.32$, $p < .001$), learning in AVLT (fifth essay–first essay) ($F[2, 89] = 19.06$, $p < .001$), delayed recall in AVLT ($F[2, 89] = 48.35$, $p < .001$), TEM1 ($F[2, 89] = 48.77$, $p < .001$), TEM2 ($F[2, 89] = 24.31$, $p < .001$), TEM3 ($F[2, 89] = 46.03$, $p < .001$), TEM4 ($F[2, 89] = 61.12$, $p < .001$), TEM5 ($F[2, 89] = 38.01$, $p < .001$) and MOCA ($F[2, 89] = 95.77$, $p < .001$).

There were no differences between groups in age, IADL, STAI_Y2 and GDS-15. Significant pairwise *post hoc* comparisons between groups are displayed in the right half of Table 1.

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

Table 2. Descriptive statistics for time intervals across tasks and groups

	Group	Mean	Std. Deviation
SVF_0–30 s	HC	8.35	1.34
	SCD	8.52	1.68
	MCI	5.50	.99
SVF_31–60 s	HC	5.18	1.64
	SCD	4.87	2.28
	MCI	3.22	1.04
SVF_61–90 s	HC	3.97	1.11
	SCD	4.19	2.04
	MCI	2.83	1.01
PVF_0–30 s	HC	5.61	2.59
	SCD	6.52	2.44
	MCI	2.03	.63
PVF_31–60 s	HC	3.81	2.19
	SCD	4.61	2.45
	MCI	1.17	.59
PVF_61–90 s	HC	2.56	1.90
	SCD	3.48	1.83
	MCI	.73	.61
Alternating_VF 0–30 s	HC	4.16	2.50
	SCD	3.58	1.18
	MCI	1.40	.50
Alternating_VF 31–60 s	HC	3.45	1.52
	SCD	2.48	.96
	MCI	1.10	.55
Alternating_VF 61–90 s	HC	2.84	1.44
	SCD	1.97	1.02
	MCI	.57	.50
OrthographicC 0–30 s	HC	5.71	2.18
	SCD	5.00	1.93
	MCI	2.50	.82
OrthographicC 31–60 s	HC	4.29	1.01
	SCD	3.65	1.17
	MCI	1.67	.66
OrthographicC 61–90 s	HC	3.52	1.15
	SCD	2.74	1.50
	MCI	1.30	.84

Descriptive statistics for time intervals across tasks and diagnostic groups are shown in Table 2.

Participants' years of education were strongly associated with all the dependent variables in the study (see Table 3); thus, this sociodemographic variable was included as a covariate in subsequent analysis.

A multivariate mixed ANOVA 3 x 4 x 3 was conducted involving one between-subjects factor (diagnostic Group), two repeated measures factors (Task and Time Interval) and the participants' years of education as a covariate. The results of this analysis pointed out that there were significant differences due to Group ($F[2, 88] = 46.82$, $p < .001$; $\eta^2_p = .52$, power = 1.00), Task ($F[3, 86] = 32.75$, $p < .001$; $\eta^2_p = .53$, power = 1.00) and Time Interval ($F[2, 87] = 18.43$, $p < .001$; $\eta^2_p = .30$, power = 1.00). There were also three significant interactions: Group x Task ($F[6, 174] = 6.21$, $p < .001$; $\eta^2_p = .17$, power = .99), Group x Time Interval ($F[4, 176] = 15.17$, $p = .012$; $\eta^2_p = .07$, power = .83) and Task x Time Interval ($F[6, 83] = 14.86$, $p < .001$; $\eta^2_p = .52$, power = 1.00.

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

Table 3. Pearson correlations between all the dependent variables and participants' years of education

	Years of education
SVF_0–30 s	.578**
SVF_31–60 s	.558**
SVF_61–90 s	.433**
PVF_0–30 s	.748**
PVF_31–60 s	.746**
PVF_61–90 s	.668**
Alternating_VF 0–30 s	.637**
Alterating_VF 31–60 s	.561**
Alternating_VF 61–90 s	.479**
OrthographicC 0–30 s	.568**
OrthographicC 31–60 s	.511**
OrthographicC 61–90 s	.212*

Note: **$p < .0001$; *$p < .05$.

Regarding the influence of the covariate, this analysis revealed that participants' years of education have an overall effect on VF performance ($F[1, 88] = 141.14$, $p < .001$; $\eta^2p = .61$, power = 1.00). This covariate also interacts significantly with Task ($F[3, 264] = 15.52$, $p < .001$; $\eta^2p = .15$, power = 1.00), Interval ($F[2, 176] = 16.79$, $p < .001$; $\eta^2p = .16$, power = .99), and Task x Interval ($F[12, 528] = 3.21$, $p = .004$; $\eta 2p = .03$, power = .93). A linear trend better characterized the relationship between the covariate and Task ($F[1, 88] = 30.48$, $p < .001$; $\eta^2p = .26$, power = 1.00) and between the covariate and Interval ($F[1, 88] = 26.18$, $p < .001$; $\eta^2p = .23$, power = .99). That is, as participants' education level increases, their performance across different tasks and time intervals tends to change in a predictable, linear way.

Next, a detailed analysis of the two former interactions (Group x Task and Group x Time Interval) was completed through univariate ANOVAs, first across Tasks (SVF, PVF, Alternating VF, and Orthographic constraint VF) and then across Intervals (0–30, 31–60, and 61–90) and using *post hoc* pairwise comparisons between Groups. In the first set of analyses, we used the total number of words produced in 90 seconds for each Task. Table 4 summarizes descriptives and significant differences between groups (pairwise comparisons) across tasks.

An inspection of Table 4 reveals that the HC and SCD groups differ significantly from the MCI group in all four tasks. On the other hand, there are no differences between the groups HC and SCD in SVF nor in PVF, but the number of words produced by the two groups differs significantly in the Alternating and Orthographic VF versions. Concerning the proposed gradient for the VF tasks in the HC group, results pointed out that there are significant differences between SVF and PVF

Table 4. Descriptives and significant differences between groups (pairwise comparisons) across tasks

	SVF				PVF				Alternating VF				Orthographic C VF			
	Mean (SD)	1	2	3	Mean (SD)	1	2	3	Mean (SD)	1	2	3	Mean (SD)	1	2	3
HC	17.5			**	11.9			**	10.4		*	**	13.5		*	**
	(3.2)				(6.2)				(4.8)				(3.3)			
SCD	17.6			**	14.6			**	8.0	*		**	11.4	*		**
	(4.8)				(6.3)				(2.2)				(3.3)			
MCI	11.5	**	**		3.9	**	**		3.1	**	**		5.4	**	**	
	(2.3)				(1.4)				(.83)				(1.7)			

Notes: 1: HC; 2: SCD; 3: MCI * p = .01, ** p < .001.

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

Table 5. Differences between groups (pairwise comparisons) across time intervals

	0–30 s				31–60 s				61–90 s			
	Mean (SD)	HC	SCD	MCI	Mean (SD)	HC	SCD	MCI	Mean (SD)	HC	SCD	MCI
HC	23.84 (7.3)			**	16.72 (5.1)			**	12.89 (4.3)			**
SCD	23.61 (5.2)			**	15.61 (4.9)			**	12.38 (3.8)			**
MCI	11.43 (2.2)	**	**		7.15 (1.5)	**	**		5.43 (2.1)	**	**	

** $p < .001$.

(paired-samples t test = 6.64, p < .001), and between the Alternating and the Orthographic versions (paired-samples t test = −4.59, p < .001).

Another set of analyses was conducted for the second interaction of interest (Group x Time Interval). For this purpose, we compute the overall number of words produced in each Time Interval (0–30, 31–60, and 61–90) by the three groups of participants. Table 5 shows significant differences between groups (pairwise comparisons) across time intervals.

Significant contrasts follow the same pattern across time intervals. The overall number of words produced by HC and SCD participants across time intervals is similar, but it differs significantly from that of the MCI group.

3.2. Relevant variables in subjects' classification

To determine their importance for subjects' classification, a discriminant analysis was computed using the overall number of words produced in each task as an independent variable. As the Box's M statistic was significant, the analysis was performed using group covariance matrices of the canonical discriminant functions. Table 6 summarizes the eigenvalues of the two discriminant functions obtained in the analysis using the Enter method, the percentage of variance explained by each one, as well as Wilks' Lambda and Chi-square statistics.

Results show that 80.04% of the original cases were correctly classified using these functions. The structure matrix (Table 7) shows that the Orthographic

Table 6. Results of discriminant analysis

Function	Eigenvalue	% of Variance	Wilks' Lambda	Chi-square	df	Sig.
1	1.63	80.6	.273	113.53	8	p < .001
2	.393	19.4	.718	29.00	3	p < .001

Table 7. Structure matrix

	Function	
	1	2
Orthographic C VF	.932*	−.069
Alternating VF	.771*	−.167
PVF	−.602	.295
SVF	.616*	.630*

Notes: Variables are ordered by the absolute size of correlation within the function. An asterisk (*) means that the variable remains in the discriminant equation at the final step.

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

constraint and the Alternating VF (in this order) have the highest correlations with the first function, explaining 80.6% of the total variance. Regarding the second function, PVF is the unique significant variable explaining the remaining 19.4% of the variance.

3.3. ROC analyses

ROC analysis using Orthographic constraint VF and HC as the positive success criteria (the negative state group includes SCD and MCI participants) showed an area under the curve (AUC) of 0.863, 95% CI (.756, .915), $p < .001$. This value means that the Orthographic constraint VF score has good diagnostic accuracy. Figure 1 illustrates the curve related to diagnostic accuracy.

The Kolmogorov–Smirnov (K-S) metric indicates that the model is good enough for classification since it is equal to 0.542 (maximum = 1.0), thus it has a medium quality to distinguish between the two groups. The optimal cut-off for the Orthographic constraint VF score based on the K-S metric for our sample was 8.5.

Concerning the Alternating VF score, ROC analysis using HC as the positive success criteria (the negative state group includes again SCD and MCI participants) showed an AUC of 0.819, 95% CI (.735, .903), $p < .001$. This value means that the Alternating VF score also has good diagnostic accuracy. Figure 2 illustrates the curve related to its diagnostic accuracy.

K-S is equal to .509, indicating that the model is good enough for classification, thus it has a medium quality to distinguish between the two groups. The optimal cut-off for the VF_ALTERN score based on the K-S metric for our sample was 4.5.

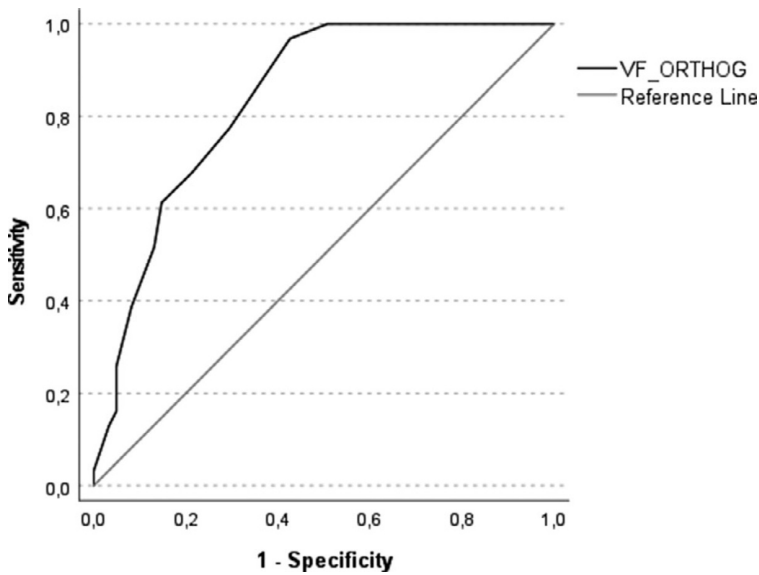


Figure 1. ROC curve for the VF with Orthographic constraint overall word productivity (VF_ORTHOG). The reference line is grey.

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

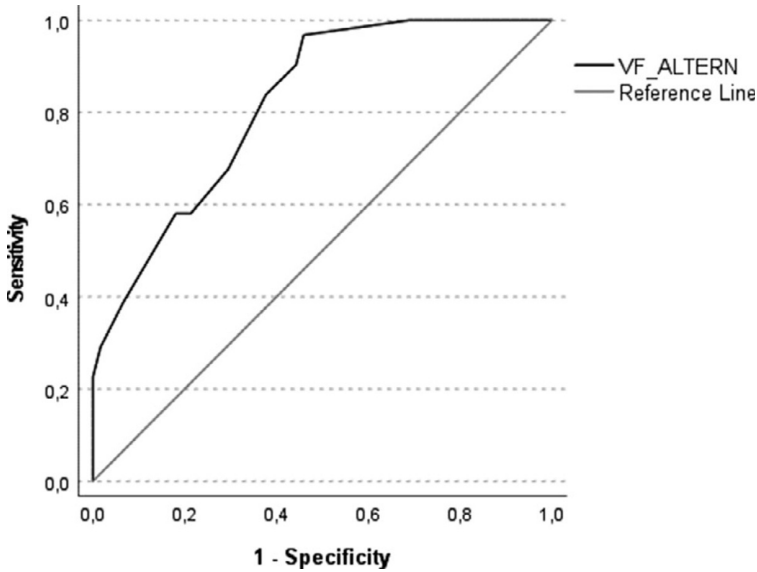


Figure 2. ROC curve for the Alternating VF overall word productivity (VF_ALTERN). The reference line is grey.

4. Discussion

4.1. Summary of findings

The present study investigated the potential benefits of extending standard VF protocols commonly used in the neuropsychological assessment of age-related cognitive decline. We hypothesized that tasks with greater executive demands – specifically, Alternating and Orthographic constrained VF – would lead to lower performance among healthy older adults (HC) and better distinguish individuals with SCD from their cognitively healthy peers. The results broadly supported these hypotheses. Among HC, they produced significantly more words in the SVF task than in the PVF task, as expected. Additionally, and somewhat unexpectedly, performance varied between the Alternating and Orthographic VF tasks, suggesting different degrees of executive demand across conditions. In terms of task difficulty, performance among HC participants followed the order: SVF > Orthographic VF > PVF > Alternating VF.

Importantly, this performance gradient was also reflected in the group comparisons. While individuals with SCD performed comparably to HC participants on the SVF and PVF tasks, they showed significantly lower performance on the more demanding VF variants. This indicates that executive-loaded tasks may be more sensitive to subtle cognitive changes and could provide added diagnostic value beyond traditional VF assessments.

As expected, MCI participants performed significantly worse than both the HC and SCD groups across all VF tasks. This is consistent with a wide range of studies documenting global cognitive deficits in MCI (McDonnell et al., 2020; Kirova et al., 2015). Nevertheless, since the MCI group had significantly fewer years of education, these differences cannot be attributed solely to cognitive impairment. Education emerged as a strong covariate across all tasks and intervals, highlighting the importance

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

of considering it as a continuous moderating factor. This underscores the need to calibrate expectations of cognitive performance according to educational background in both clinical and research settings.

Nonetheless, the findings reinforce the idea that both traditional (SVF and PVF) and executive-demanding VF tasks are effective tools for identifying cognitive decline at different stages.

4.2. Executive demands in fluency tasks

The performance pattern observed in the HC group aligns well with theoretical expectations regarding executive demand. SVF, which relies primarily on automatic semantic retrieval, was the easiest (Shores et al., 2006). In contrast, both PVF and orthographic VF required greater cognitive control, while Alternating VF imposed the highest executive demands due to the need for mental set shifting, attentional monitoring and inhibition (Lam & Marquardt, 2022; Macoir et al., 2022). This gradient is consistent with previous research (e.g., Heim et al., 2009; Macoir et al., 2022; Shao et al., 2014), reinforcing the idea that tasks imposing multiple executive constraints place a heavier burden on cognitive resources.

Notably, orthographic VF appeared less demanding than PVF for HC participants, despite involving inhibitory control. One possible explanation is that semantic-based retrieval (e.g., naming body parts) may benefit from more automatic associative processes (Macoir et al., 2022), which remain relatively preserved under orthographic constraints – especially in transparent orthographies like Portuguese. PVF, by contrast, involves artificial constraints and suppresses natural retrieval strategies, thereby increasing cognitive load (Heim et al., 2009; Marko et al, 2023).

These task-specific demands were particularly relevant for differentiating SCD from HC individuals. While previous studies have reported inconsistent findings regarding the sensitivity of SVF and PVF in early cognitive impairment (Nikolai et al., 2018; Aiello et al., 2023), our results emphasize the diagnostic value of tasks specifically designed to engage executive functions. The significant performance gap between SCD and HC groups on executive-loaded VF tasks supports the hypothesis that early executive dysfunction characterizes SCD, a view echoed by recent studies (Jessen et al., 2023; López-Higes et al., 2017; Macoir et al., 2022).

Previous studies have also reported that older adults with SCD tend to exhibit slightly lower average performance across several cognitive domains compared to those without SCD (Kielb et al., 2017; Morrison, 2023). These findings suggest that some individuals with SCD may experience subtle neuropsychological deficits that do not yet meet the clinical threshold for MCI but can still be identified through standardized neuropsychological testing. Wolfsgruber et al. (2020) found that lower performance on composite scores assessing verbal memory, executive function, language and global cognition was associated with elevated CSF biomarkers of AD in memory clinic patients with SCD. This supports the notion that *minor neuropsychological deficits* in individuals with SCD may signal emerging AD pathology. Patients with SCD who demonstrate subtle deficits on cognitive tests not only show higher levels of AD biomarkers but also face an increased risk of cognitive decline and clinical progression. Therefore, the presence of these minor neuropsychological impairments may help identify a high-risk subgroup among individuals seeking medical advice for subjective cognitive concerns (Stark et al., 2023).

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

4.3. Clinical implications

Provided that all VF tasks are brief and easy-to-administer assessment instruments, it would be highly recommended to include both classic tasks (SVF and PVF) and tasks with greater executive load (alternating and with orthographic constraint) in neuropsychological assessment protocols. The former tasks would be sensitive in more advanced stages of deterioration (MCI), whereas the latter variants have good diagnostic accuracy to detect patients with incipient executive problems (preclinical stage SCD). Both discriminant and ROC analyses confirmed the diagnostic accuracy of these tasks, with areas under the curve exceeding 0.80 and clear cut-off points (≤ 8.5 for Orthographic VF; ≤ 4.5 for Alternating VF). These benchmarks are consistent with prior findings (Macoir et al., 2022; Alegret et al., 2018) and provide clinicians with actionable thresholds for early detection. This is especially valuable in primary care or community settings where access to biomarkers or neuroimaging may be limited.

There are some previous studies that find results compatible with this line of reasoning. For example, the study conducted by Alegret et al. (2018) with a Spanish sample aimed to explore the presence of VF deficits in individuals with MCI and mild AD dementia. The researchers also sought to evaluate the utility of VF tests in identifying cognitively healthy individuals who are at risk of progressing to MCI, as well as those with MCI who may later develop dementia. Additionally, the study aimed to establish specific VF cut-off scores that could be useful for cognitive assessment within the Spanish population. To achieve these objectives, the study involved a large sample comprising 568 cognitively healthy participants, 885 individuals with MCI and 367 with mild AD dementia. All participants underwent VF testing alongside a comprehensive neuropsychological battery. Furthermore, longitudinal analyses were conducted on a subset of the sample (231 cognitively healthy individuals and 667 MCI subjects) to identify whether VF performance could predict future diagnosis conversions over time. The results demonstrated that lower performance on VF tests was significantly associated with the progression from cognitively healthy status to MCI, as well as from MCI to dementia. When analyzing the influence of time until diagnosis conversion, VF performance showed a significant effect in predicting faster progression from normality to MCI. However, this predictive effect was not observed for the transition from MCI to dementia. The researchers also calculated VF cut-off scores along with sensitivity and specificity values for six different conditions, which were stratified by three age ranges and two educational levels. In conclusion, their findings suggest that VF testing is a valuable neuropsychological tool for detecting different stages of cognitive decline related to AD. Since VF deficits appear early in the disease process, this assessment can be particularly useful not only for identifying cognitively healthy individuals at risk of developing MCI but also for monitoring the progression from MCI to dementia.

McDonnell et al. (2020) evaluate the ability of SVF (animals) and PVF (FAS) to discriminate among normal aging, amnestic-MCI (a-MCI) and AD. Their study included a sample of 332 participants with a mean age equal to 65.75 (99 healthy older adults, 90 a-MCI and 43 AD patients). Results indicate that SVF and PVF were significant predictors of diagnostic classification, but semantic fluency explained a greater amount of the discriminant ability of the model. These results suggest from a practical standpoint that, particularly, SVF may be an accurate and efficient tool in screening early dementia in time-limited clinical/medical settings. This research also

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

supported the fact that VF, especially SVF, declines progressively in AD, with semantic fluency deteriorating faster than phonemic fluency as the disease advances.

There is only one previous study showing that VF tasks with higher executive load are useful for detecting cognitive deficits in the preclinical stage of AD. Macoir et al. (2022) conducted a research in which a total of 60 adults with SCD and 60 HCs, all of them French-speaking Canadians, performed one free action (verb) fluency task, an alternating fluency task (between words beginning with the letter T and words belonging to the clothes category) and an orthographic constraint fluency task (names of animals whose written form did not involve the letter A). The performance of the participants with SCD and the HCs in the free action fluency task was similar. However, HCs performed significantly better than SCD in the alternating fluency task, which involved cognitive flexibility, and the orthographic constraint fluency task, which required inhibition.

5. Conclusions and limitations of the study

In HC, task performance revealed a clear gradient of difficulty aligned with increasing executive demands, ranging from superior performance in SVF to progressively lower output in the Orthographic constraint, PVF and Alternating fluency tasks.

While participants with SCD matched HCs in traditional VF tasks (SVF and PVF), their reduced performance in the more demanding tasks – particularly those requiring cognitive flexibility and inhibition – supports the hypothesis that VF tasks with higher executive load are sensitive indicators of early cognitive changes in the preclinical phase of AD. These findings align with prior research emphasizing early executive dysfunction in SCD and reinforce the diagnostic value of complex VF paradigms.

Results obtained in the study also highlight the need to consider education not as a categorical variable but as a continuous covariate in cognitive assessment. This approach enhances the precision and fairness of neuropsychological evaluations.

Finally, from a clinical point of view, the study advocates for the integration of both traditional and executive-loaded VF tasks into routine neuropsychological assessments. While classic SVF and PVF tasks remain useful for identifying cognitive decline in more advanced stages (e.g., MCI), tasks such as Alternating VF and Orthographic constraint VF offer significant diagnostic accuracy for detecting early executive deficits, particularly in SCD. Establishing task-specific cut-off scores further contributes to the development of reliable and efficient screening tools for early detection and monitoring of AD-related cognitive decline.

In sum, VF tasks – especially those incorporating higher executive demands – should be regarded as essential components in the early identification and longitudinal monitoring of neurodegenerative processes.

It should be noted as a limitation of the study that results are obtained from a small sample of participants in each group, which limits generalizability. The SCD sample shows some degree of heterogeneity in terms of gender, age and educational background; however, its representativeness of the broader population of cognitively healthy individuals with subjective cognitive complaints is limited. The recruitment strategy, although designed to access individuals with varied cognitive and sociodemographic profiles, relied on convenience sampling. These institutions are more likely to attract older adults who are socially engaged, potentially more health conscious and possibly better functioning than the general population of older adults

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

with subjective cognitive complaints; thus, this issue should be considered when assessing the scope of the results obtained (Snitz et al., 2018).

The cross-sectional design used is also a limitation that should be overcome in future longitudinal studies, given that SCD with more attentional/executive and linguistic difficulties seems to have a higher risk of AD (Valech et al., 2017). Larger and homogeneous clinical cohorts are also specific goals for future studies. Additionally, future analyses could benefit from incorporating qualitative metrics such as clustering and switching, which may offer further insight into the specific cognitive mechanisms affected in preclinical stages of dementia.

Data availability statement. The data and syntax code describing all statistical analyses, and the detailed description of the data structure (included in the data file), are accessible at: https://osf.io/tbrgm/?view_only=344126762cb24431947fbe540b4c424d

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

Competing interests. The authors declare none.

References

Aiello, E. N., Preti, A. N., Pucci, V., Diana, L., Corvaglia, A., Barattieri di San Pietro, C., … Bolognini, N. (2022). The Italian telephone-based verbal fluency battery (t-VFB): Standardization and preliminary clinical usability evidence. *Frontiers in Psychology*, 13, 963164. https://doi.org/10.3389/fpsyg.2022.963164.

Aiello, E. N., Verde, F., Solca, F., Milone, I., Giacopuzzi Grigoli, E., Dubini, A., … Poletti, B. (2023). Lower semantic fluency scores and a phonemic-over-semantic advantage predict abnormal CSF P-tau181 levels in aβ+ patients within the Alzheimer's disease clinical spectrum. *Neurological Sciences*, 44 (6), 1979–1985. https://doi.org/10.1007/s10072-023-06643-w.

Aita, S. L., Beach, J. D., Taylor, S. E., Borgogna, N. C., Harrell, M. N., & Hill, B. D. (2019). Executive, language, or both? An examination of the construct validity of verbal fluency measures. *Applied Neuropsychology: Adult*, 26 (5), 441–451. https://doi.org/10.1080/23279095.2018.1439830.

Alegret, M., Peretó, M., Pérez, A., Valero, S., Espinosa, A., Ortega, G., … Boada, M. (2018). The role of verb fluency in the detection of early cognitive impairment in Alzheimer's disease. *Journal of Alzheimer's Disease*, 62 (2), 611–619. https://doi.org/10.3233/JAD-170826.

Apostle, J. L. A., de Jesus Loureiro, L. M., dos Reis, I. A. C., da Silva, I. A. L. L., Cardoso, D. F. B., & Sfetcu, R. (2014). Contribution to the adaptation of the geriatric depression Scale-15 into Portuguese. *Nursing Journal Reference*, 4 (3), 65–73. https://doi.org/10.12707/RIV14033.

Araújo, L., Ribeiro, O., & Paúl, C. (2017). The role of existential beliefs within the relation of centenarians' health and well-being. *Journal of Religion and Health*, 56, 1111–1122. https://doi.org/10.1007/s10943-016-0297-5.

Bairami, S., Folia, V., Liampas, I., Ntanasi, E., Patrikelis, P., Siokas, V., … Kosmidis, M. H. (2023). Exploring verbal fluency strategies among individuals with normal cognition, amnestic and non-amnestic mild cognitive impairment, and Alzheimer's disease. *Medicina*, 59 (10), 1860. https://doi.org/10.3390/medicina59101860.

Baldo, J. V., Schwartz, S., Wilkins, D. P., & Dronkers, N. F. (2010). Double dissociation of letter and category fluency following left frontal and temporal lobe lesions. *Aphasiology*, 24 (12), 1593–1604. https://doi.org/10.1080/02687038.2010.489260.

Bieling, P. J., Antony, M. M., & Swinson, R. P. (1998). The state--trait anxiety inventory, trait version: Structure and content re-examined. *Behaviour Research and Therapy*, 36 (7–8), 777–788. https://doi.org/10.1016/s0005-7967(98)00023-0.

Cavaco, S., Goncalves, A., Pinto, C., Almeida, E., Gomes, F., Moreira, I., … Teixeira-Pinto, A. (2013). Semantic fluency and phonemic fluency: Regression-based norms for the Portuguese population. *Archives of Clinical Neuropsychology*, 28 (3), 262–271. https://doi.org/10.1093/arclin/act001.

Cavaco, S., Gonçalves, A., Pinto, C., Almeida, E., Gomes, F., Moreira, I., … Teixeira-Pinto, A. (2015). Auditory verbal learning test in a large nonclinical Portuguese population. *Applied Neuropsychology: Adult*, 22 (5), 321–331. https://doi.org/10.1080/23279095.2014.927767.

Cipolotti, L., MacPherson, S. E., Gharooni, S., van-Harskamp, N., Shallice, T., Chan, E., & Nachev, P. (2018). Cognitive estimation: Performance of patients with focal frontal and posterior lesions. *Neuropsychologia*, 115, 70–77. https://doi.org/10.1016/j.neuropsychologia.2017.08.017.

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 (8), 954–962. https://doi.org/10.1093/arclin/acw071.

Costa, A., Bagoj, E., Monaco, M., Zabberoni, S., De Rosa, S., Papantonio, A. M., … 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, 365–372. https://doi.org/10.1007/s10072-013-1520-8.

Delgado-Losada, M. L., Rubio-Valdehita, S., López-Higes, R., Campos-Magdaleno, M., Avila-Villanueva, M., Frades-Payo, B., & Lojo-Seoane, C. (2024). Phonological fluency norms for Spanish middle-aged and older adults provided by the SCAND initiative (P, M, & R). *Journal of the International Neuropsychological Society*, 30(2), 172–182. https://doi.org/10.1017/S1355617723000309.

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.

Dowling, N. M., Hermann, B., La Rue, A., & Sager, M. A. (2010). Latent structure and factorial invariance of a neuropsychological test battery for the study of preclinical Alzheimer's disease. *Neuropsychology*, 24 (6), 742. https://doi.org/10.1037/a0020176.

Ekström, I., Josefsson, M., Bäckman, L., & Laukka, E. J. (2024). Predictors of cognitive aging profiles over 15 years: A longitudinal population-based study. *Psychology and Aging*, 39 (5), 467–483. https://doi.org/10.1037/pag0000807.

Fernandes, S. M., Araújo, A. M., Vázquez-Justo, E., Pereira, C., Silva, A., Paul, N., … Maestú, F. (2018). Effects of aging on memory strategies: A validation of the Portuguese version of the test of memory strategies. *The Clinical Neuropsychologist*, 32 (1), 133–151. https://doi.org/10.1080/13854046.2018.1490456.

Freitas, S., Simões, M. R., Martins, C., Vilar, M., & Santana, I. (2010). Estudos de adaptação do Montreal cognitive assessment (MoCA) Para a população Portuguesa. *Avaliaçao Psicologica: Interamerican Journal of Psychological Assessment*, 9(3), 345–357.

García-González, S. (2024). Can phonemic verbal fluency be used to predict Alzheimer's disease? *NeuroSci*, 5 (4), 501–508. https://doi.org/10.3390/neurosci5040036.

Graf, C. (2008). The Lawton instrumental activities of daily living scale. *AJN The American Journal of Nursing*, 108 (4), 52–62. https://doi.org/10.1097/01.NAJ.0000314810.46029.74.

Heim, S., Eickhoff, S. B., Friederici, A. D., & Amunts, K. (2009). Left cytoarchitectonic area 44 supports selection in the mental lexicon during language production. *Brain Structure and Function*, 213 (4), 441–456. https://doi.org/10.1007/s00429-009-0213-9.

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia*, 42 (9), 1212–1222. https://doi.org/10.1016/j.neuropsychologia.2004.02.001.

Hirnstein, M., Stuebs, J., Moè, A., & Hausmann, M. (2023). Sex/gender differences in verbal fluency and verbal-episodic memory: A meta-analysis. *Perspectives on Psychological Science*, 18 (1), 67–90. https://doi.org/10.1177/17456916221082116.

Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., … Silverberg, N. (2018). NIA-AA research framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia*, 14 (4), 535–562. https://doi.org/10.1016/j.jalz.2018.02.018.

Jessen, F., Amariglio, R. E., Van Boxtel, M., Breteler, M., Ceccaldi, M., Chételat, G., … Subjective Cognitive Decline Initiative. (2014). A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimer's & Dementia*, 10 (6), 844–852. https://doi.org/10.1016/j.jalz.2014.01.001.

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

Jessen, F., Wolfsgruber, S., Kleineindam, L., Spottke, A., Altenstein, S., Bartels, C., … Düzel, E. (2023). Subjective cognitive decline and stage 2 of Alzheimer disease in patients from memory centers. *Alzheimer's & Dementia*, 19 (2), 487–497. https://doi.org/10.1002/alz.12674.

Julian, L. J. (2011). Measures of anxiety. *Arthritis Care & Research*, 63(11), 10–1002. https://doi.org/10.1002/acr.20561.

Kielb, S., Rogalski, E., Weintraub, S., & Rademaker, A. (2017). Objective features of subjective cognitive decline in a United States national database. *Alzheimer's Dement.*, 13 (12), 1337–1344. https://doi.org/10.1016/j.jalz.2017.04.0088.

Kiselica, A. M., Webber, T. A., & Benge, J. F. (2020). The uniform dataset 3.0 neuropsychological battery: Factor structure, invariance testing, and demographically adjusted factor score calculation. *Journal of the International Neuropsychological Society*, 26 (6), 576–586. https://doi.org/10.1017/S135561772000003X.

Kirova, A. M., Bays, R. B., & Lagalwar, S. (2015). Working memory and executive function decline across normal aging, mild cognitive impairment, and Alzheimer's disease. *BioMed Research International*, 2015 (1), 748212. https://doi.org/10.1155/2015/748212.

Koppara, A., Wagner, M., Lange, C., Ernst, A., Wiese, B., König, H. H., … Jessen, F. (2015). Cognitive performance before and after the onset of subjective cognitive decline in old age. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 1 (2), 194–205. https://doi.org/10.1016/j.dadm.2015.02.005.

Lam, B. P., & Marquardt, T. P. (2022). Factors predicting mental effort associated with verbal fluency: Cue types, switching, and fear of negative evaluation. *Journal of Speech, Language, and Hearing Research*, 65 (9), 3491–3501. https://doi.org/10.1044/2022_JSLHR-21-0056.

Lawton, M. P. (1969). Assessment of older people; self-maintaining and instrumental activities of daily living. *Gerontologist*, 9, 279–286. https://doi.org/10.1093/geront/9.3_Part_1.179.

Linz, N., Fors, K. L., Lindsay, H., Eckerström, M., Alexandersson, J., & Kokkinakis, D. (2019). Temporal analysis of the semantic verbal fluency task in persons with subjective and mild cognitive impairment. In *Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology* (pp. 103–113). https://doi.org/10.18653/v1/W19-3012.

López-Higes, R., Prados, J. M., Rubio, S., Montejo, P., & Del Río, D. (2017). Executive functions and linguistic performance in SCD older adults and healthy controls. *Aging, Neuropsychology, and Cognition*, 24 (6), 717–734. https://doi.org/10.1080/13825585.2016.1256370.

López-Higes, R., Rubio-Valdehita, S., Fernández-Blázquez, M. A., Lojo-Seoane, C., Ávila-Villanueva, M., Montenegro-Peña, M., … 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.

Lubrini, G., Periáñez, J. A., Laseca-Zaballa, G., Bernabéu-Brotons, 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.

Macoir, J., Tremblay, P., & Hudon, C. (2022). The use of executive fluency tasks to detect cognitive impairment in individuals with subjective cognitive decline. *Behavioral Sciences*, 12 (12), 491. https://doi.org/10.3390/bs12120491.

Marko, M., Michalko, D., Dragašek, J., Vančová, Z., Jarčušková, D., & Riečanský, I. (2023). Assessment of automatic and controlled retrieval using verbal fluency tasks. *Assessment*, 30 (7), 2198–2211. https://doi.org/10.1177/10731911221117512.

McDonnell, M., Dill, L., Panos, S., Amano, S., Brown, W., Giurgius, S., … Miller, K. (2020). Verbal fluency as a screening tool for mild cognitive impairment. *International Psychogeriatrics*, 32 (9), 1055–1062. https://doi.org/10.1017/S1041610219000644.

Morrison, C. O. (2023). Subjective cognitive decline is associated with lower baseline cognition and increased rate of cognitive decline. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, 78 (4), 573–584. https://doi.org/10.1093/geronb/gbac1789.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The Montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53 (4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x.

Nevado, A., Del Rio, D., Martin-Aragoneses, M. T., Prados, J. M., & Lopez-Higes, R. (2021). Preserved semantic categorical organization in mild cognitive impairment: A network analysis of verbal fluency. *Neuropsychologia*, 157, 107875. https://doi.org/10.1016/j.neuropsychologia.2021.107875.

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

McDonough, I. M., & Allen, R. S. (2019). Biological markers of aging and mental health: A seed and soil model of neurocognitive disorders. *Aging & Mental Health*, 23 (7), 793–799. https://doi.org/10.1080/13607863.2018.1531383.

Nikolai, T., Bezdicek, O., Markova, H., Stepankova, H., Michalec, J., Kopecek, M., … Vyhnalek, M. (2018). Semantic verbal fluency impairment is detectable in patients with subjective cognitive decline. *Applied Neuropsychology: Adult*, 25 (5), 448–457. https://doi.org/10.1080/23279095.2017.1326047.

Núcleo de Estudos de Geriatria da Sociedade Portuguesa de Medicina Interna – GERMI (2017). *Avaliação Geriátrica* [Internet]. Available from: https://www.spmi.pt/docs_nucleos/GERMI_36.pdf.

Pekkala, S. (2012). Verbal fluency tasks and the neuropsychology of language. *The handbook of the neuropsychology of language*, 1, 619–634. https://doi.org/10.1002/9781118432501.

Peña-Casanova, J., Quinones-Ubeda, S., Gramunt-Fombuena, N., Quintana-Aparicio, M., Aguilar, M., Badenes, D., … NEURONORMA Study Team. (2009). 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.

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, 256 (3), 183–194. https://doi.org/10.1111/j.1365-2796.2004.01388.x.

Petersen, R. C. (2011). Mild cognitive impairment. *New England Journal of Medicine*, 364 (23), 2227–2234. https://doi.org/10.1056/NEJMcp0910237.

Petersen, R. C., Caracciolo, B., Brayne, C., Gauthier, S., Jelic, V., & Fratiglioni, L. (2014). Mild cognitive impairment: A concept in evolution. *Journal of Internal Medicine*, 275 (3), 214–228. https://doi.org/10.1111/joim.12190.

Piatt, A. L., Fields, J. A., Paolo, A. M., & Tröster, A. I. (2004). Action verbal fluency normative data for the elderly. *Brain and Language*, 89 (3), 580–583. https://doi.org/10.1016/j.bandl.2004.02.003.

Rami, L., Molinuevo, J. L., Sanchez-Valle, R., Bosch, B., & Villar, A. (2007). Screening for amnestic mild cognitive impairment and early Alzheimer's disease with M@ T (memory alteration test) in the primary care population. *International Journal of Geriatric Psychiatry: A journal of the psychiatry of late life and allied sciences*, 22 (4), 294–304. https://doi.org/10.1002/gps.1672.

Rodrigues, P. F. S., Bártolo, A., Ribeiro, B., López-Higes, R., Rubio-Valdehita, S., Caetano, A. P., & Fernandes, S. M. (2025). Everyday memory questionnaire-revised (EMQ-R): Psychometric validation of the European Portuguese version in non-clinical sample. *Behav Sci (Basel)*, 15 (3), 280. https://doi.org/10.3390/bs15030280.

Royle, J., & Lincoln, N. B. (2008). The everyday memory questionnaire–revised: Development of a 13-item scale. *Disability and Rehabilitation*, 30 (2), 114–121. https://doi.org/10.1080/09638280701223876.

Ryan, J. J., & Geisser, M. E. (1986). Validity and diagnostic accuracy of an alternate form of the Rey auditory verbal learning test. *Archives of Clinical Neuropsychology*, 1 (3), 209–217. https://doi.org/10.1016/0887-6177(86)90027-2.

Sanabria, A., Alegret, M., Rodriguez-Gomez, O., Valero, S., Sotolongo-Grau, O., Monté-Rubio, G., … Boada, M. (2018). The Spanish version of face-name associative memory exam (S-FNAME) performance is related to amyloid burden in subjective cognitive decline. *Scientific Reports*, 8 (1), 3828. https://doi.org/10.1038/s41598-018-21644-y.

Santos Nogueira, D., Azevedo Reis, E., & Vieira, A. (2017). Verbal fluency tasks: Effects of age, gender, and education. *Folia Phoniatrica et Logopaedica*, 68 (3), 124–133. https://doi.org/10.1159/000450640.

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. A., Carstairs, J. R., & Crawford, J. R. (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.

Silva, D. R., & Spielberger, C. D. (2007). Manual do inventário de estado-traço de ansiedade (STAI). *Mind Garden*, 62.

Snitz, B. E., Wang, T., Cloonan, Y. K., Jacobsen, E., Chang, C. C. H., Hughes, T. F., … Ganguli, M. (2018). Risk of progression from subjective cognitive decline to mild cognitive impairment: The role of study setting. *Alzheimer's & dementia: the journal of the Alzheimer's Association*, 14 (6), 734. https://doi.org/10.1016/j.jalz.2017.12.003.

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press

Spielberger, C. D., Gonzalez-Reigosa, F., Martinez-Urrutia, A., Natalicio, L. F., & Natalicio, D. S. (1971). The state-trait anxiety inventory. *Revista Interamericana de Psicologia/Interamerican journal of psychology*, 5(3 & 4). https://doi.org/10.30849/rip/ijp.v5i3%20&%204.620.

Stark, M., Wolfsgruber, S., Kleineidam, L., Frommann, I., Altenstein, S., Bartels, C., … Wagner, M. (2023). Relevance of minor neuropsychological deficits in patients with subjective cognitive decline. *Neurology*, 101 (21), e2185–e2196. https://doi.org/10.1212/WNL.0000000000207844.

Troyer, A. K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology*, 22 (3), 370–378. https://doi.org/10.1076/1380-3395(200006)22:3;1-V;FT370.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology*, 11 (1), 138. https://doi.org/10.1037/0894-4105.11.1.138.

Valech, N., Tort-Merino, A., Coll-Padrós, N., Olives, J., León, M., Rami, L., & Molinuevo, J. L. (2017). Executive and language subjective cognitive decline complaints discriminate preclinical Alzheimer's disease from normal aging. *Journal of Alzheimer's Disease*, 61 (2), 689–703. https://doi.org/10.3233/JAD-170627.

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.

Vicente, S. G., Benito-Sánchez, I., Barbosa, F., Gaspar, N., Dores, A. R., Rivera, D., & Arango-Lasprilla, J. C. (2022). Normative data for verbal fluency and object naming tests in a sample of European Portuguese adult population. *Applied Neuropsychology: Adult*, 29 (5), 1268–1279. https://doi.org/10.1080/23279095.2020.1868472.

Whiteside, D. M., Kealey, T., Semla, M., Luu, H., Rice, L., Basso, M. R., & Roper, B. (2016). Verbal fluency: Language or executive function measure? *Applied Neuropsychology: Adult*, 23 (1), 29–34. https://doi.org/10.1080/23279095.2015.1004574.

Wolfsgruber, S., Kleineidam, L., Guski, J., Polcher, A., Frommann, I., Roeske, S., … DELCODE Study Group. (2020). Minor neuropsychological deficits in patients with subjective cognitive decline. *Neurology*, 95 (9), e1134–e1143. https://doi.org/10.1212/WNL.0000000000010142.

Wright, L. M., De Marco, M., & Venneri, A. (2023). Verbal fluency discrepancies as a marker of the prehippocampal stages of Alzheimer's disease. *Neuropsychology*, 37 (7), 790. https://doi.org/10.2147/PRBM.S284645.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17 (1), 37–49. https://doi.org/10.1016/0022-3956(82)90033-4.

Yubero, R., Gil, P., Paul, N., & Maestú, F. (2011). Influence of memory strategies on memory test performance: A study in healthy and pathological aging. *Aging, Neuropsychology, and Cognition*, 18 (5), 497–515. https://doi.org/10.1080/13825585.2011.597840.

Cite this article: López-Higes, R., Fernandes, S. M., Rodrigues, P. F. S., Paiva, B., Rubio-Valdehita, S., & Del Río, D. (2025). Verbal fluency with high executive load for early dementia's risk detection, *Language and Cognition*, 17, e76, 1–22. https://doi.org/10.1017/langcog.2025.10033

https://doi.org/10.1017/langcog.2025.10033 Published online by Cambridge University Press