Version of Record: https://www.sciencedirect.com/science/article/pii/S0278262623000179
Manuscript_7dc3a242de8f6b7f0381cf55fde5d8ea

Gray Matter Volume and Within-task Verbal Fluency Performance Among Older Adults

Brigitte Pace[A]

Roee Holtzer[A,B]

Mark E. Wagshul[C,D]

[A] Ferkauf Graduate School of Psychology, Yeshiva University; 1165 Morris Park Ave, The Bronx, NY 10461

[B] Department of Neurology, Albert Einstein College of Medicine, 1300 Morris Park Ave, The Bronx, NY 10416

[C] Department of Radiology, Gruss Magnetic Resonance Research Center, Albert Einstein College of Medicine; 1250 Morris Park Ave, The Bronx, NY 10461

[D] Physiology and Biophysics, Albert Einstein College of Medicine; 1300 Morris Park Ave, The Bronx, NY 10416

bpace@mail.yu.edu (corresponding author)

Roee.holtzer@einsteinmed.edu

Mark.wagshul@einsteinmed.edu

Abstract

The current study examined the relationship between gray matter volume (GMV) and rate of word generation over the course of three consecutive 20-sec intervals in 60-sec letter and category verbal fluency (VF) tasks. Attenuated rate of within-person word generation in VF provides incremental information beyond total scores and predicts increased risk of incident Mild Cognitive Impairment (MCI). No studies to date, however, have determined the structural neural substrates underlying word generation rate in VF. Participants were 70 community-residing adults ≥ 65 years, who completed the letter and category VF tasks and a 3T structural MRI scan. Linear mixed effects models (LMEMs) were used to determine the moderating effect of GMV on word generation rate. Whole brain voxel-wise LMEMs, adjusted for age, gender, education, Wide-Range Achievement Test – reading subtest score (WRAT3), and global health score, were run using permutation methods to correct for multiple comparisons. Lower GMV, primarily in frontal regions (superior frontal, rostral middle frontal, frontal pole, medial orbitofrontal, and pars orbitalis), were related to attenuated word generation rate, especially for letter VF. We propose that lower frontal GMV underlies inefficient executive word search processes reflected by attenuated word generation slope in letter VF amongst older adults.

KEYWORDS: word generation rate, magnetic resonance imaging (MRI), healthy older adults, category verbal fluency, letter verbal fluency, executive functions

1.1 Introduction

Verbal Fluency (VF) is commonly used in neuropsychological assessments and requires individuals to generate words that belong to a specific category (e.g., animals/fruits/vegetables) or begin with a specific letter (e.g., F/A/S) over a fixed time period, typically 60-sec. Word generation in VF involves semantic memory, verbal abilities (McDowd et al., 2011), sustained attention, working memory, cognitive flexibility (Diamond, 2013), speed of processing (Bryan, Luszcz, & Crawford, 1997), and executive processes such as initiation of word retrieval (Henry et al., 2004; Monsch et al., 1994), utilization of search strategies, monitoring responses, and inhibiting intrusions (Henry et al., 2004). Traditionally, the total number of criterion-conforming words generated is scored. Total words can be used to differentiate various patient and non-patient groups (e.g., Sung et al., 2012), particularly when it is interpreted in the context of an array of neuropsychological test data. However, total word score provides limited insight into the complex underlying processes involved.

Word generation rate slows as individuals progress through the 60-sec task (Fernaeus, et al., 2008; Fernaeus & Almkvist, 1998; Raboutet et al., 2010), with more rapid generation at the outset and slower generation as time passes (Crowe, 1998). One study demonstrated that almost 50% of the total words generated in 60-sec VF tasks were produced within the first 15-sec (Venegas & Mansur, 2011). This reduction in word generation rate is thought to reflect changes in the underlying neurocognitive processes involved. Previous research identified this change in word generation rate and labeled it “intraindividual variability” in VF, measured by within-person word generation slope during task performance (Holtzer et al., 2020). Various definitions of intraindividual variability, both conceptually and operationally, assert that it is a measurable feature of performance that reflects the within-person change in the neurocognitive underpinnings involved (MacDonald et al., 2006). Generally, intraindividual variability has been identified as a useful indicator of integrity of the nervous system (Hultsch et al., 2000) and it provides incremental information beyond measures of central tendency (Dixon et al., 2007). Intraindividual variability among neuropsychological tasks is related to neurocognitive impairment (Holtzer et al., 2008) and variability within measures is related to compromised executive control processes (West et al., 2002) associated with frontal areas (Stuss et al., 1999). Intraindividual variability in VF provides information beyond total scores (Fernaeus et al., 2008) and is linked to mild cognitive impairment (MCI; Demetriou & Holtzer, 2017). Greater within-person word generation slopes (i.e., steeper negative slope), attributed to higher word generation during the initial and more automatic phase of word retrieval, predicted lower risk of incident MCI (Holtzer et al., 2020).

An understanding of more nuanced scoring, such as within-person word generation slopes, here referred to as VF slope, can further our understanding of the cognitive processes involved in letter versus category VF. Both types of VF rely on verbal memory (Moscovitch, 1994) and the temporal and frontal lobes, but each type also has distinct cognitive and neural correlates. Category VF (CVF) uses more semantic processes, relying on left temporal areas (Murphy et al., 2006), while letter VF (LVF) uses more strategic lexical search strategies (Henry et al., 2004), relying on left frontal areas (Li et al., 2017). However, the cognitive mechanisms underlying each task are not easily delineated. Factor analyses have shown mixed results: in one study, LVF loaded on executive functioning/speed factors while CVF loaded on language factors (Kiselica et al., 2020); in others, LVF (e.g., Dowling et al., 2010) and CVF loaded on the language factor with limited support for an executive contribution to either type (Whiteside et al., 2015). In yet another study, both LVF and CVF loaded onto an “executive” factor (Aita et al., 2019). Importantly, verbal fluency is multidimensional in nature and maps onto cognitive domains including language, executive functions, and processing speed, making it difficult to delineate specific neurocognitive contributions to task performance. Understanding VF slopes in cognitively normal aging individuals could provide additional information regarding the nuances of the underlying neurocognitive mechanisms of VF.

Neuroimaging evidence identifies specific brain regions and networks that correlate with neuropsychological task performance, promoting our understanding of brain-behavior relationships. This research provides some support for the notion that VF is an executively driven task (e.g., Pietrzykowski et al., 2021; Yuan & Raz, 2014). Neuroimaging evidence is often derived from patient or lesion-based studies and indicates that widely distributed brain regions are implicated in VF total word scores. Given that VF relies at least in part on executive functioning (Gustavson et al., 2019; Lezak, 1995) housed in the frontal lobes, it is unsurprising that lesions in the dorsolateral prefrontal cortex result in VF impairment (e.g., bilateral superior medial and dorsomedial frontal areas; Stuss et al., 1998). Left frontal lesions (Baldo & Shimamura, 1998; Baldo et al., 2001; Robinson et al., 2012) are associated with both types of VF; lesions *and* activation in parietal and thalamic regions are associated with CVF and LVF (Birn et al. 2010; Stuss et al., 1998; Wagner et al., 2014). The semantic component of VF is thought to be reflected by temporal cortical involvement (e.g., Ghanavati et al., 2019). Functional imaging studies have found similar results suggesting that LVF and CVF are associated with prefrontal and temporal regions, such as blood oxygen level dependent responses in the perisylvian language network (Curtis et al., 2001). Cortical areas unique to CVF have been identified in left temporal regions via association with temporal lobe degradation (Troyer et al., 1998; Henry & Crawford, 2004), lesions (Baldo et al., 2006; Biesbroek et al., 2015), and activation (Birn et al., 2010). Activation in the right inferior frontal gyrus (Buckner et al., 1995; Watanabe et al., 1998; Dan et al., 2013;) is also uniquely associated with CVF. Those unique to LVF include left posterior and dorsal inferior frontal gyrus (Costafreda et al., 2006; Heim et al., 2009; Biesbroek et al., 2015) and supplementary motor area (Schlösser et al., 1998; Cook et al., 2014).

Notably, most existing research does not include associations between gray matter volume (GMV) and VF performance in healthy aging. In one study of community-dwelling adults aged ≥ 65, inferior frontal GMV was positively associated with early (0-15-sec) VF performance whereas left hippocampal GMV was positively associated with late (16-30-sec) performance in one out of four categories on a 30-sec CVF task (Catheline et al., 2015). No study, to our knowledge, has investigated the association of GMV with the *change* in word generation rate (i.e., slope) during 60-sec VF tasks in healthy older adults. Of particular note, the heterogeneity in the extant literature concerning associations between brain integrity and VF task performance highlights the challenges involved in identifying specific neural substrates of complex neurocognitive processes.

1.2 Current study

We examined the relationship between GMV and LVF and CVF slope. VF slope was defined as the changes in word generation from 0-60-sec, with word count taken at three time intervals (0-20-sec [T1], 21-40-sec [T2], 41-60-sec [T3]). Whole-brain GMV analyses were carried out, using the FreeSurfer image analysis suite (documented and available at http://surfer.nmr.mgh.harvard.edu). First, we established that, as expected, word generation declined within each task. Next, we examined the relationship between GMV and total word scores. Finally, we hypothesized that lower GMV would be associated with attenuated slopes in both tasks. A data-driven approach using whole-cortex data was used given the insufficient evidence to predict specific regions associated with VF slope. However, given that VF is associated with frontal regions, frontal GMV is particularly likely to be implicated in the interaction between GMV and VF slope. It was expected that there would be both shared and distinct regions of GMV implicated in slope in CVF compared to LVF.

2.1 Methods

70 right-handed community-residing older adults (M_{age}=74.9 ±5.3, 50% female; see Table 1) were enrolled in Central Control of Mobility in Aging, a larger study examining cognitive and brain imaging predictors of mobility. Potential participants were identified from publicly available population data of lower Westchester County, New York. Structured phone interviews were used to obtain verbal consent and determine eligibility including dementia screens (Galvin et al., 2005; Galvin, Roe, Xiong, & Morris, 2006; Lipton et al., 2003), medical history, and physical functioning. General exclusion criteria: inability to speak/understand English, inability to walk independently, residence in a nursing home, visual/auditory loss, diagnosis of a serious/acute illness, psychiatric condition, or neurodegenerative disease, and presence of a neurological gait disorder. The main study assessed mobility, (instrumental) activities of daily living, and psychological, motoric, and social functioning. Additional standard magnetic resonance imaging (MRI) exclusion criteria (surgically implanted metal devices, claustrophobia, etc.) were used to select participants for the MRI sub-study.

Cognitive status was confirmed through case conference procedures (Holtzer et al., 2008) carried out as part of the research protocol implemented in the Central Control of Mobility in Aging Study (Holtzer, Wang, & Verghese, 2014). The study was HIPPA-compliant, approved by the Institutional Review Board of Albert Einstein College of Medicine, and completed in accordance with Helsinki Declaration. Detailed study procedures are outlined elsewhere (Holtzer, Wang, & Verghese, 2014).

2.2 Outcome Measures. VF was measured using The Controlled Oral Word Association Test (COWAT; Spreen & Benton, 1977), including three 60-sec letter (F, A, S) and three category (Fruits, Vegetables, Animals) fluencies. LVF was administered before CVF. Time intervals ranging from 10 to 30-sec have been used in the literature to examine the 60-sec task (Fernaeus & Almkvist, 1998; Raboutet et al., 2010; Weakley et al., 2013). Previous work (Demetriou & Holtzer, 2017) demonstrated that 20-sec intervals are an appropriate dissection of the task, allowing for more than two data points in the word generation curve while also allowing for a sufficient tally of words generated during each interval. Therefore, correct words at three consecutive 20-sec intervals (T1, T2, T3) were recorded for each iteration of the task without changing the standard administration of the task. Participants were instructed to avoid responses that included proper nouns or similar responses with different suffixes (e.g., *fix, fixing*). Errors were determined according to the following: words that did not fit in the identified category or that started with a letter other than the target letter, proper nouns, and repetitions, including similar words with different suffixes. Only correct responses were used to calculate scores. Within each VF type, the number of words generated was summed across trials for each interval and used to derive individual word generation slopes.

2.3 Covariates. Global Health Status (GHS) was computed using a dichotomous rating indicating the presence or absence of 10 chronic conditions (diabetes, chronic heart failure, arthritis, hypertension, depression, stroke, Parkinson’s disease, chronic obstructive lung disease, angina, and myocardial infarction) with scores ranging from 0-10 (Holtzer et al., 2008). The Wide-Range Achievement Test (WRAT3) – Reading Subtest (Wilkinson, 1993) was administered to measure premorbid intellectual function; it requires participants to pronounce printed words increasing in complexity and unfamiliarity. The WRAT3 has been extensively used and validated in the aging literature and was thus included in the current study.

2.4 Magnetic resonance imaging. MRI was performed on a 3T Philips scanner (Achieva TX, Philips Medical Systems, Best, The Netherlands) with a 32-channel head coil. A single high-resolution, T1-weighted image was used for all analyses (MPRAGE – TE/TR/TI = 4.6/9.8/900ms, voxel size 1mm isotropic, Sense acceleration factor 2.6).

2.5 Data processing and analysis. Preprocessing for MRI was performed with the Freesurfer image analysis suite. Freesurfer's morphometric procedures (see Fischl et al., 2004; Salat et al., 2004) have good test-retest reliability across scanner manufacturers and field strengths (Reuter et al., 2012). The processing uses a hybrid watershed/surface formation procedure to remove non-brain tissue (Ségonne et al., 2004), automated Talairach transformation, segmentation of the subcortical white matter and deep gray matter volumetric structures (Fischl et al., 2002), intensity normalization (Sled et al., 1998), tessellation of the gray-white boundary, automated topology correction (Fischl et al., 2001), and surface deformation following intensity gradients to optimally identify tissue class transitions (Dale et al., 1999; Fischl & Dale, 2000). After creating the cortical model, cerebral cortex was parceled into sub-units according to gyral and sulcal structure (Desikan et al., 2006; Fischl et al., 2004). Maps of the cortical surface were created using spatial intensity gradients across tissue classes and are thus not merely reliant on absolute signal intensity. Volumes were calculated through Freesurfer's recon-all pipeline, across whole-brain cortical and subcortical grey matter structures.

Multiple steps were taken to ensure optimal accuracy in mapping subjects' images during processing. The quality of each participant's automatic parcellation data ("APARC") was checked manually by overlaying the participants' cortical segmentation on their original T1 images. No adjustments to automated cortical surface estimations were needed. Individual cortical maps were projected onto the Freesurfer standard space ("fsaverage") and smoothed with a 5mm Gaussian kernel to allow for variation across subjects in cortical features and registration to the standard.

2.6 Statistical analysis and calculation. Fully adjusted linear mixed-effects models (LMEMs; see below) were carried out using a modified version of the Freesurfer LMEM software written for Matlab. In contrast to the standard Freesurfer LMEM model which enters cortical volume as the dependent variable, our modified version considered letter or category generation rate as the dependent variable and brain volume on a voxel-wise basis as an independent variable. Volume and other covariates were transformed to subject-wise standardized Z-scores. All models reported included time and task as main effects and were adjusted for age, gender, education, GHS, and raw WRAT score. Two-way interactions (Time x GMV) determined if the effect of time on slope was moderated by regional GMV; time was coded as 1, 2, and 3 for the three 20-sec VF periods (T1, T2, T3) as a continuous variable so that only the linear VF slopes were considered. Separate general linear models (GLMs) were run to examine the areas associated with total word scores.

Cortical vertices across the whole brain were entered into the statistical models. Permutation Analysis of Linear Models (PALM), a software that allows inferences and correction for multiple comparisons across the cortex, was used to analyze the processed dataset with 2,000 permutations using the Freedman-Lane procedure (Winkler et al., 2014). Threshold-free cluster enhancement (TFCE) identified clusters of vertices significant at the 0.05 level, avoiding the need to establish an arbitrary cluster-defining threshold; evidence suggests that

TFCE provides better sensitivity and more interpretable output than cluster-based thresholding (Smith & Nichols, 2009).

SPSS statistical software package, (version 25; SPSS, Inc., Chicago, IL) was used for statistical analysis.

The fully adjusted LMEMs considered time-by-volume as an interaction term, time as a three-level repeated within-subject continuous variable, demographic and behavioral covariates as fixed effects, and participant as random effects (see equation).

Mean word generation per time interval ~	…outcome
Time*Volume	…interaction
Time (Interval)	…main effect
Volume	…main effect
Age+Gender+Education+WRAT3+GHS	…covariates
(1\Subject)	…random effects

2.7 Data Visualization. To visually represent the interaction effects across the cortex, regions with significant effects from the LMEM were converted into a color scale based on beta estimates and projected onto the standard fsaverage cortical surface.

3.1 Results

Participants included 70 older adults (M_{age} = 74.9 ±5.3 years; $M_{\text{education}}$ = 15.5 ±3.2 years; 50% female). The sample was relatively healthy (GHS M=1.3 ±1.04) with average premorbid function ($M_{\text{WRAT-3}}$ Standard Score = 108.9 ±9.7 see Table 1). As predicted and in line with existing literature, word generation rate decreased during CVF and LVF (see Table 2), as demonstrated by the main effect of time in LVF and CVF (Betas ranged from -4.63 to -4.50 and -4.60 to -4.58 words/interval respectively, p values < 0.001). GLMs considering total word count and covariates revealed no significant areas of GMV associated with total word generation in either VF type.

[INSERT TABLE 1 HERE]

[INSERT TABLE 2 HERE]

Main effects and covariates were included in all models that examined the time-by-GMV interaction (see supplemental tables). As hypothesized, beta estimates for the interactions using whole-brain data were negative, indicating that *lower* GMV in identified areas was associated with *attenuated* slopes in both tasks (see Table 3). There were no significant regions with higher GMV associated with attenuated slopes (LVF $p > 0.17$; CVF $p > 0.61$).

[INSERT TABLE 3 HERE]

3.2 LVF: Lower volume in the following frontal, parietal, and occipital areas was associated with attenuated LVF slope: left caudal middle frontal, left lateral orbitofrontal, left pars orbitalis, left rostral middle frontal, left frontal pole, left medial orbitofrontal, left superior frontal, left superior parietal, bilateral isthmus cingulate, bilateral precuneus, bilateral cuneus, bilateral lateral occipital, bilateral lingual, and bilateral pericalcarine (see Figure 1 and Table 3).

3.3 CVF: Lower volume in the following left hemisphere frontal and parietal areas was associated with attenuated slope in CVF: paracentral, precentral, postcentral, and posterior cingulate (see Table 3 and Figure 1, for right hemisphere $p > 0.057$).

[INSERT FIGURE 1 HERE]

3.4 LVF and CVF Shared Regions. There was small to no meaningful overlap between only two regions (left hemisphere parietal) with GMV significantly related to *both* types of fluency. Notably, regions within the left precuneus appear to have some role in both types of VF, however, the amount of cortex associated with LVF is markedly larger than that associated with CVF slope (21.71% and < 1% of total left precuneus volume respectively). The overlap in the left superior parietal region also has a larger amount of cortical volume related to LVF than to CVF slope (2.24% and < 1% of total left superior parietal respectively).

3.5 Main effect of volume. LMEMs revealed significant main effects of traditional sum scores (i.e., total score over 60-sec) in CVF in left hemisphere paracentral (β = 2.490 words/cc, 23.6% of total left paracentral volume, $p = 0.048$), left hemisphere posterior cingulate (β = 2.746 words/cc, 10.1% of total left posterior cingulate volume, $p = 0.046$), and right hemisphere pars opercularis (β = 3.244 words/cc, 34.9% of total right pars opercularis volume, $p = 0.043$). No statistically significant main effects of traditional sum score were found for LVF.

4.1 Discussion

The current study was designed to determine GMV correlates of CVF and LVF slopes, or slopes of within-task word generation rate, in community-dwelling older adults. Consistent with previous research (Crowe, 1998; Fernaeus & Almkvist, 1998; Raboutet et al., 2010), word generation slopes were negative, demonstrating a decline in word generation over the course of the 60-sec tasks. As evidenced via significant brain region volume by time interactions demonstrated by LMEMs, results offer compelling evidence that GMV moderated the decline in VF slope over time. Reduced volume in select cortical areas was associated with *attenuated* VF slopes primarily, but not exclusively, in LVF. These findings align with existing research reporting attenuated VF slopes are associated with poor outcomes (i.e., MCI and incident MCI; Demetriou & Holtzer, 2017; Holtzer et al., 2020) and are interpreted to suggest that individuals with reduced negative slopes are less able to maximize the early processes of word retrieval involved in rapidly accessing the lexico-semantic store. Producing fewer words during the first interval results in having less to lose at later intervals when more effortful search strategies are employed. Current results suggest that individuals with greater GMV in select frontal and posterior regions are better able to utilize the rapid automatic retrieval from the easily accessed pool of words at the outset of the task. Later, these individuals then have a steeper reduction in word generation rate.

Stuss's (1998) study used 15-sec intervals to demonstrate a decrease in word generation rate among patients with frontal grey matter lesions. Current results expand on this work in three ways: first, we demonstrate the importance of frontal grey matter for word generation in a healthy older adult population, second, we identified the specific frontal grey matter regions driving these effects, and third, we quantified the relationship between reduced GMV and attenuated word generation rate in these regions. Current findings also expand upon Catheline et al.'s (2015) work, by identifying areas associated with the *change* in word generation between three 20-sec intervals across the 60-sec VF task, rather than the areas associated with 0-15-sec and 16-30-sec word generation.

Fully adjusted models demonstrated that cortical regions related to VF slope, primarily LVF, included left hemisphere frontal areas. Most of the areas implicated were related to LVF but not CVF slopes, this may reflect that LVF slope maps onto the cortex better than CVF slope or that LVF is more executively driven than CVF; the latter interpretation is supported by prior studies (Henry et al., 2004; Costafreda et al., 2006; Heim et al., 2009; Biesbroek et al., 2015; Schlösser et al., 1998; Cook et al., 2014). Current results, implicating left frontal regions, are consistent with conceptualizations of intraindividual variability as a distinct executive process subserved, at least in part, by the frontal cortex (Holtzer et al., 2020). Areas implicated are key for executive functions including attentional control, working memory coordination, and inhibition, and align with evidence that LVF is primarily executively and frontally driven (e.g., Henry et al., 2004; Li et al., 2017). These results also support the notion that LVF slope is associated with specific executive functions in this version of the task.

The left hemisphere caudal middle frontal (Hanford et al., 2019), lateral orbitofrontal (Ghahremani et al., 2010; Hampshire et al., 2012; Watson et al., 2018), and pars orbitalis (Aron et al., 2014; Hartwigsen et al., 2019) were implicated in LVF slope and are each associated with attention. The implication of attention in LVF slope might suggest that initiation of LVF search strategies requires cognitive activation and initiation of an attentionally mediated search strategy and that attentional control is required to inhibit inappropriate strategies or responses. The left rostral middle frontal (Grambaite et al., 2011; Nakamura-Palacios et al., 2014) and left medial orbitofrontal cortices (mOFC; Szatowska et al., 2007) are involved in inhibition. Caudal middle frontal cortex (Marqués-Iturria et al., 2014), left superior frontal gyrus (Cutini et al., 2008), and left rostral middle frontal cortex (Grambaite et al., 2011) are associated with switching. Given the association of these regions with inhibition and switching and our data involving a novel measure (slope) of an executive task reliant on alphabetical search strategies (LVF), results that were not reproduced when using total word scores, offer compelling evidence that VF slope reflects these features of executive search strategies.

Executive functions are often intertwined; for example, the medial orbitofrontal cortex (mOFC) and the lateral orbitofrontal cortex (lOFC) are both associated with attention and switching. Left mOFC is implicated in complex processes combining attentional switching and response inhibition, such as inhibiting prepotent responses in favor of less habitual ones (Szatowska et al., 2007) exerting top-down control on decisional impulsive behavior (Wang et al., 2019). Similarly, the lOFC is associated with attentional control (Hampshire & Owen, 2006) and inhibition when a dominant/habitual response needs to be overridden by a weaker but goal-relevant response (Ghahremani et al., 2010; Hampshire et al., 2012; Watson et al., 2018).

The overlapping roles of lOFC and left mOFC in marshalling attentional control and inhibiting habitual/impulsive responses to prioritize novel, goal-oriented ones suggest that they involve shared neurocognitive and executive functions. The association of LVF slope with these regions likely reflects implementation of novel search strategies, relying heavily on adapting strategies or increasing cognitive control during the task. Switching attention to less well-learned response sets (i.e., alphabetic search strategies) over more habitual ones (i.e., semantic search strategies) is necessary in LVF and may be a key component in successful early performance. Monitoring conflicts between competing search strategies and attentional control (i.e., directing attention away from inappropriate/ineffective strategies) are linked to working memory and other executive processes associated with VF (Aita et al., 2019).

Working memory is related to frontal areas identified in the current results. Left frontal pole (Brodmann's Area 10) GMV is associated with multi-tasking (Gilbert et al., 2006), monitoring and integrating subgoals during working memory tasks (Braver & Bongiolatti, 2002), programming articulatory motor movements (Faulkner & Wilshire, 2020), and multi-tasking language skills (i.e., translating between languages simultaneously in bilingual individuals; Becker et al., 2016). The left superior frontal gyrus plays a specific role in the coordination of working memory (Alagapan et al., 2019; du Boisgueheneuc et al., 2006); and blood oxygenation levels in this region increase during task-switching (Cutini et al., 2008). Thus, results suggest that verbal working memory and switching ability are implicated in LVF slope.

CVF slope was generally related to areas of GMV that are associated with the motor component of speech production (Itabashi et al., 2016), language (i.e., speech initiation and fluency; Dick et al., 2019), and attention and executive functioning (Habets et a., 2019; Zhang et al., 2015; Coderre & van Heuven, 2013; i.e., left precentral and postcentral areas). This may reflect that CVF recruits neural circuitry of more semantically routinized information, involving recitation and visualization. CVF slope was also associated with volume in parietal regions associated with performance on executive tasks in older adults (Respino et al., 2019) including.

shifting performance (Roye et al., 2020) and attentional focus (Leech & Sharp, 2014; Hahn et al., 2007), as well as internally directed cognition (Buckner et al., 2008). The involvement of these regions (i.e., left postcentral and posterior cingulate) in CVF slope suggests that attentional control is involved in semantic search processes reflected by word generation rate.

Results indicate a small to insignificant overlap between regions of GMV associated with LVF vs. CVF slopes. This aligns with evidence that different neurocognitive processes underlie the two tasks, with CVF more semantically driven (Murphy et al., 2006) and LVF more reliant on strategic lexical search strategies (Henry et al., 2004). Ultimately, results add to the line of reasoning that, although the brain substrates involved in *both* tasks are likely executive in nature, the manner in which executive functioning is involved evidently varies across task type; that is, the strategies through which these processes are employed throughout task performance vary meaningfully.

In both LVF and CVF, the involvement of semantic and executive processes changes from beginning to end of the task (Fernaeus & Almkvist, 1998; Crowe, 1998; Raboutet et al., 2010). Initial word generation relies on low-effort, rapid access to a lexico-semantic stock, thus individuals can retrieve more readily available and frequently used words; later word generation relies on executive search strategies and retrieval of lower frequency words (Raboutet et al., 2010; Lezak et al., 2012; Crowe, 1998; Fernaeus & Almkvist, 1998). Change in word generation rate appears to be a useful indicator of changes among these processes. Indeed, the inflection point (first inter-word delay >2 standard deviations than the preceding inter-word intervals; Barois et al., 2020) is thought to represent the shift from rapid access to the semantic stock into more executive search strategies. While the operationalization of VF slope herein precludes the estimation of inflection time, the two may be related. That is, Barois et al.'s (2020) inflection time and the current study's VF slope may tap into similar neurocognitive processes. The extensive frontal GMV involvement, particularly regarding changes in early word generation rate, indicates an executively driven (i.e., attentional control, inhibition of habitual processes, etc.) transition between cognitive processes during VF performance.

4.2 Limitations, strengths, and future directions

We examined change in word generation rate using three consecutive 20-sec intervals in relation to GMV. LMEMs were run on a whole-brain voxel-wise basis, including demographic and behavioral covariates. Given that word generation rate is correlated with time and total word generation, it is challenging to interpret the main effects of volume and time on word generation. It would therefore not be appropriate to compare current LMEMs to previous studies examining structural brain correlates of total word generation, which did not include the effect of interval, or time, and their interaction. Our separate GLM did not reveal significant regions of GMV in association with total word generation. Given that structural correlates of VF slope were the focus of the current study, our GLM was likely underpowered to identify this relationship in a healthy cohort of 70 adults. This relationship should be further investigated; indeed, some existing literature begins to address this line of inquiry (e.g., Catheline et al., 2015).

The current study's specified LMEMs enabled a focused investigation of word generation rate and GMV across the cortex. The repeated measurements within a person (i.e., intervals) increased the power to identify GMV correlates of VF slopes. Permutation testing corrected for multiple comparisons, allowing the thousands of LMEMs to maintain a family-wise alpha level of 0.05. This process of controlling for multiple comparisons is considerably more conservative than other methods of correction and contributes to the interpretation that the results offer compelling evidence for the role of frontal involvement in VF slope, particularly in LVF (e.g., Puoliväli et al., 2020).

Results indicate insignificant overlap between regions in which GMV differences influence LVF vs. CVF slopes; there are several important caveats to consider. First, current models adjusted for reading scores and education (WRAT3). The variability in WRAT3 score and education level was relatively restricted within our sample; individuals scored primarily in the Average to High Average range on the WRAT3 and earned ≥12 years of education. Measures of intelligence are highly associated with education level, such as Weschler Information subtests, and have been shown to be more strongly associated with category than letter fluency in young samples (e.g., Ardila et al., 2000). Thus, while the current results align with evidence that different neurocognitive processes underlie the two tasks, the investigation of VF slope should be extended to a more diverse sample in terms of intellectual functioning and educational attainment and investigated with and without adjusting for these factors. In addition, research suggests an age-dependent contribution of cognitive skills to each type of VF performance (Stolwyk et al., 2015). Specifically, semantic retrieval was associated with CVF in younger adults, but not in older adults; verbal intelligence and processing speed were associated with LVF in younger adults, but only verbal intelligence was associated with LVF in older adults. It is possible that CVF slope would have additional clear neural associates in a younger sample. It is important to consider the possible influence of neural dedifferentiation, an age-related phenomenon in which brain functions that are typically localized to specific brain areas in younger adults are less localized in people of more advanced age (e.g., Koen et al., 2020). Thus, the cortical areas associated with LVF slope in our healthy aging sample may be different or more pronounced in a younger adult sample.

Future research should explore VF slope in relation to cortical thickness and should consider utilizing machine learning technology; these avenues of inquiry and methodology may reveal additional or more nuanced brain-behavior relationships in VF performance. Additional research should also further elucidate the role of neural substrates in VF slope with larger and more diverse samples. Examining neuroimaging correlates among different diseases associated with changes in cognition and brain structure would illuminate the neural substrates involved. Additionally, as the field of clinical neuropsychology seeks to become more comprehensive across cultures and languages, the incorporation of languages other than English, non-alphabetic languages, and individuals from diverse cultural backgrounds is essential in cultivating a more complete understanding of this widely used task. VF is associated with executive functioning in individuals who speak non-alphabetic languages; neural substrates associated with language production are similar in alphabetic and non-alphabetic speakers (Tao et al., 2020).

4.3 Conclusion.

Current results provide initial evidence that VF word generation slopes tap into a unique aspect of performance in aging individuals. Results suggest that left frontal regions including the caudal middle frontal, lateral orbitofrontal, pars orbitalis, rostral middle frontal, frontal pole, medial orbitofrontal, and the superior frontal areas are involved in LVF word generation rate. This aligns with extant literature suggesting that cognitive search strategies, particularly in LVF, are frontally and executively driven. Future work should continue the pursuit of understanding underlying neurocognitive processes involved in VF slope and differences in LVF versus CVF slope in cognitively normal and disease populations.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships, which may be considered as potential competing interests:

This work was supported by the National Institute on Aging (R01AG036921-01A1). The funding source had no involvement in the writing of the report nor the decision to submit this article for publication.

References

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

Alagapan, S., Lustenberger, C., Hadar, E., Shin, H. W., & Fröhlich, F. (2019). Low-frequency direct cortical stimulation of left superior frontal gyrus enhances working memory performance. *NeuroImage*, *184*, 697–706. https://doi.org/10.1016/j.neuroimage.2018.09.064

Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists*, *15*(1), 31–36.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: one decade on. *Trends in cognitive sciences*, *18*(4), 177–185. https://doi-org.elibrary.einsteinmed.org/10.1016/j.tics.2013.12.003

Baldo, J. V., & Shimamura, A. P. (1998). Letter and category fluency in patients with frontal lobe lesions. *Neuropsychology*, *12*(2), 259–267. https://doi.org/10.1037//0894-4105.12.2.259

Baldo, J. V., Shimamura, A. P., Delis, D. C., Kramer, J., & Kaplan, E. (2001). Verbal and design fluency in patients with frontal lobe lesions. *Journal of the International Neuropsychological Society : JINS*, *7*(5), 586–596. https://doi.org/10.1017/s1355617701755063

Baldo, J. V., Schwartz, S., Wilkins, D., & Dronkers, N. F. (2006). Role of frontal versus temporal cortex in verbal fluency as revealed by voxel-based lesion symptom mapping. *Journal of the International Neuropsychological Society : JINS*, *12*(6), 896–900. https://doi.org/10.1017/S1355617706061078

Barois, E., Sagawa, Y., Yilmaz, S., Magnin, E., & Decavel, P. (2020) What (more) can verbal fluency tell us about multiple sclerosis? *Ann Phys Rehabil Med. 2020*;S1877-0657(20)30108-1. doi:10.1016/j.rehab.2020.05.002

Becker, M., Schubert, T., Strobach, T., Gallinat, J., & Kühn, S. (2016). Simultaneous interpreters vs. professional multilingual controls: Group differences in cognitive control as well as brain structure and function. *NeuroImage*, *134*, 250–260. https://doi.org/10.1016/j.neuroimage.2016.03.079

Biesbroek, J. M., van Zandvoort, M. J. E., Kappelle, L. J., Velthuis, B. K., Biessels, G. J., & Postma, A. (2015). Shared and distinct anatomical correlates of semantic and phonemic fluency revealed by lesion-symptom mapping in patients with ischemic stroke. *Brain Struct. Funct(221)* 2123-2134. doi: 10.1007/s00429-015-1033-8

Birn, R. M., Kenworthy, L., Case, L., Caravella, R., Jones, T. B., Bandettini, P. A., & Martin, A. (2010). Neural systems supporting lexical search guided by letter and semantic category cues: a self-paced overt response fMRI study of verbal fluency. *NeuroImage*, *49*(1), 1099–1107. https://doi.org/10.1016/j.neuroimage.2009.07.036

Braver, T. S., & Bongiolatti, S. R. (2002). The role of frontopolar cortex in subgoal processing during working memory. *NeuroImage*, *15*(3), 523–536. https://doi.org/10.1006/nimg.2001.1019

Bryan, J. Luszcz, M. A., & Crawford, J. R. (1997) Verbal knowledge and speed of information processing as mediators of age differences in verbal fluency performance among older adults. *Psychology and Aging. 12(3),* 473-478. 0882-7974/97/S3.00.

Buckner, R. L., Raichle, M. E., & Petersen, S. E. (1995). Dissociation of human prefrontal cortical areas across different speech production tasks and gender groups. J. *Neurophysiol. 74*, 2163–2173. https://doi.org/10.1152/jn.1995.74.5.2163

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*, 1–38. https://doi.org/10.1196/annals.1440.011

Catheline, G., Amieva, H., Dilharreguy, B., Bernard, C., Duperron, M. G., Helmer, C., . . . Allard, M. (2015). Semantic retrieval over time in the aging brain: Structural evidence of hippocampal contribution. *Hippocampus, 25*(9), 1008-1016. doi:10.1002/hipo.22423

Coderre, E. L., & van Heuven, W. J. (2013). Modulations of the executive control network by stimulus onset asynchrony in a Stroop task. *BMC neuroscience*, *14*, 79. https://doi.org/10.1186/1471-2202-14-79

Cook, P. A., McMillan, C. T., Avants, B. B., Peelle, J. E., Gee, J. C., & Grossman, M. (2014). Relating brain anatomy and cognitive ability using a multivariate multimodal framework. *Neuroimage 99*, 477–486. https://doi.org/10.1016/j.neuroimage.2014.05.008

Costafreda, S. G., Fu, C. H. Y., Lee, L., Everitt, B., Brammer, M. J., & David, A. S. (2006). A systematic review and quantitative appraisal of fMRI studies of verbal fluency: role of the left inferior frontal gyrus. *Hum. Brain Mapp. 27*(10), 799–810. https://doi.org/10.1002/hbm.20221

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*, 391-401.

Curtis, V. A., Dixon, T. A., Morris, R. G., Bullmore, E. T., Brammer, M. J., Williams, S. C., … McGuire, P. K. (2001). Differential frontal activation in schizophrenia and bipolar illness during verbal fluency. *J. Affect. Disord., 66*, 111-121. https://doi.org/10.1016/S0165-0327(00)00240-8

Cutini, S., Scatturin, P., Menon, E., Bisiacchi, P. S., Gamberini, L., Zorzi, M., & Dell'Acqua, R. (2008). Selective activation of the superior frontal gyrus in task-switching: an event-related fNIRS study. *NeuroImage*, *42*(2), 945–955. https://doi.org/10.1016/j.neuroimage.2008.05.013

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9 (2), 179-194. https://doi.org/10.1006/nimg.1998.0395.

Dan, H., Dan, I., Sano, T., Kyutoku, Y., Oguro, K., Yokota, H., … Watanabe, E., (2013). Language-specific cortical activation patterns for verbal fluency tasks in Japanese as assessed by multichannel functional near-infrared spectroscopy. *Brain Lang. 126*(2), 208–216. https://doi.org/10.1016/j.bandl.2013.05.007

Demetriou, E. & Holtzer, R. (2017). Mild Cognitive Impairments Moderate the Effect of Time on Verbal Fluency Performance. *J Int Neuropsychol Soc, 23*(1), 44-55. doi:10.1017/S1355617716000825.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., … Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage, 31(3),* 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Dick, A. S., Garic, D., Graziano, P., & Tremblay, P. (2019). The frontal aslant tract (FAT) and its role in speech, language and executive function. *Cortex; a journal devoted to the study of the nervous system and behavior*, *111*, 148–163. https://doi.org/10.1016/j.cortex.2018.10.015

Dixon, R. A., Garrett, D. D., Lentz, T. L., MacDonald, S. W., Strauss, E., & Hultsch, D. F. (2007). Neurocognitive markers of cognitive impairment: Exploring the roles of speed and inconsistency. *Neuropsychology*, 21, 381–399. http://dx.doi.org/10.1037/0894-4105.21.3.381

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), 742–756. https://doi.org/10.1037/a0020176

du Boisgueheneuc, F., Levy, R., Volle, E., Seassau, M., Duffau, H., Kinkingnehun, S., …Dubois, B. (2006). Functions of the left superior frontal gyrus in humans: a lesion study. *Brain : a journal of neurology*, *129*(Pt 12), 3315–3328. https://doi.org/10.1093/brain/awl244

Faulkner, J. W., & Wilshire, C. E. (2020). Mapping eloquent cortex: A voxel-based lesion-symptom mapping study of core speech production capacities in brain tumour patients. *Brain and language*, *200*, 104710. https://doi.org/10.1016/j.bandl.2019.104710

Fernaeus, S.-E, & Almkvist, O. (1998). Word production: Dissociation of two retrieval modes of semantic memory across time. *Journal of Clinical and Experimental Neuropsychology, 20,* 137-143.

Fernaeus, S. E., Ostberg, P., Hellström, A., & Wahlund, L. O. (2008). Cut the coda: early fluency intervals predict diagnoses. *Cortex; a journal devoted to the study of the nervous system and behavior*, *44*(2), 161–169. https://doi.org/10.1016/j.cortex.2006.04.002

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. Natl. Acad. Sci. U. S. A. 97 (20), 11050-11055. https://doi.org/10.1073/pnas.200033797.

Fischl, B., Liu, A., & Dale, A. M. (2001). Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE transactions on medical imaging*, *20*(1), 70–80. https://doi.org/10.1109/42.906426

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., … Dale, A. M. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron, 33(3),* 341–355. https://doi.org/10.1016/s0896-6273(02)00569-x

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., … Dale, A. M. (2004). Automatically parceling the human cerebral cortex. Cerebr. Cortex 14 (1), 11-22. http://doi.org/10.1093/cercor/bhg087.

Galvin, J. E., Roe, C. M., Powlishta, K. K., Coats, M. A., Muich, S. J., Grant, E., ... Morris, J. C. (2005). The AD8: A brief informant interview to detect dementia. Neurology, 65, 559 – 564. http://dx.doi.org/ 10.1212/01.wnl.0000172958.95282.2a

Galvin, J. E., Roe, C. M., Xiong, C., & Morris, J. C. (2006). Validity and reliability of the AD8 informant interview in dementia. Neurology, 67, 1942–1948.

Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D., & Burgess, P. W. (2006). Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis. *Journal of cognitive neuroscience*, *18*(6), 932–948. https://doi.org/10.1162/jocn.2006.18.6.932

Ghanavati, E., Salehinejad, M. A., Nejati, V., & Nitsche, M. A. (2019). Differential role of prefrontal, temporal and parietal cortices in verbal and figural fluency: Implications for the supramodal contribution of executive functions. *Scientific reports*, *9*(1), 3700. https://doi.org/10.1038/s41598-019-40273-7

Ghahremani, D. G., Monterosso, J., Jentsch, J. D., Bilder, R. M., & Poldrack, R. A. (2010). Neural components underlying behavioral flexibility in human reversal learning. *Cerebral cortex (New York, N.Y. : 1991)*, *20*(8), 1843–1852. https://doi.org/10.1093/cercor/bhp247

Grambaite, R., Selnes, P., Reinvang, I., Aarsland, D., Hessen, E., Gjerstad, L., & Fladby, T. (2011). Executive dysfunction in mild cognitive impairment is associated with changes in frontal and cingulate white matter tracts. *Journal of Alzheimer's disease: JAD*, *27*(2), 453–462. https://doi.org/10.3233/JAD-2011-110290

Grogan, A., Green, D. W., Ali, N., Crinion, J. T., & Price, C. J. (2009). Structural correlates of semantic and phonemic fluency ability in first and second languages. *Cerebral cortex (New York, N.Y. : 1991)*, *19*(11), 2690–2698. https://doi.org/10.1093/cercor/bhp023

Gustavson, D. E., Panizzon, M. S., Franz, C. E., Reynolds, C. A., Corley, R. P., Hewitt, J. K., … & Friedman, N. P. (2019). Integrating verbal fluency with executive functions: Evidence from twin studies in adolescence and middle age. *Journal of experimental psychology. General*, *148*(12), 2104–2119. https://doi.org/10.1037/xge0000589

Habets, E., Hendriks, E. J., Taphoorn, M., Douw, L., Zwinderman, A. H., Vandertop, W. P., … Klein, M. (2019). Association between tumor location and neurocognitive functioning using tumor localization maps. *Journal of neuro-oncology*, *144*(3), 573–582. https://doi.org/10.1007/s11060-019-03259-z

Hahn, B., Ross, T. J., & Stein, E. A. (2007). Cingulate activation increases dynamically with response speed under stimulus unpredictability. *Cerebral cortex (New York, N.Y. : 1991)*, *17*(7), 1664–1671. https://doi.org/10.1093/cercor/bhl075

Hampshire, A., Chaudhry, A. M., Owen, A. M., & Roberts, A. C. (2012). Dissociable roles for lateral orbitofrontal cortex and lateral prefrontal cortex during preference driven reversal learning. *NeuroImage*, *59*(4), 4102–4112. https://doi.org/10.1016/j.neuroimage.2011.10.072

Hampshire, A., & Owen, A. M. (2006). Fractionating attentional control using event-related fMRI. *Cerebral cortex (New York, N.Y.: 1991)*, *16*(12), 1679–1689. https://doi.org/10.1093/cercor/bhj116

Hanford, L. C., Pinnock, F., Hall, G. B., & Heinrichs, R. W. (2019). Cortical thickness correlates of cognitive performance in cognitively-matched individuals with and without schizophrenia. *Brain and cognition*, *132*, 129–137. https://doi.org/10.1016/j.bandc.2019.04.003

Hartwigsen, G., Neef, N. E., Camilleri, J. A., Margulies, D. S., & Eickhoff, S. B. (2019). Functional Segregation of the Right Inferior Frontal Gyrus: Evidence From Coactivation-Based Parcellation. *Cerebral cortex (New York, N.Y. : 1991)*, *29*(4), 1532–1546. https://doi.org/10.1093/cercor/bhy049

Heim, S., Eickhoff, S. B., & Amunts, K. (2009). Different roles of cytoarchitectonic BA 44 and BA 45 in phonological and semantic verbal fluency as revealed by dynamic causal modelling. *Neuroimage 48*(3), 616–624. https://doi.org/10.1016/j.neuroimage.2009.06.044

Henry, J. D., & Crawford, J. R. (2004). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology*, *18*(2), 284–295. https://doi.org/10.1037/0894-4105.18.2.284

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. doi:10.1016/j.neuropsychologia.2004.02.001

Holtzer, R., Jacobs, S., & Demetriou, E. (2020). Intraindividual variability in verbal fluency performance is moderated by and predictive of mild cognitive impairments. *Neuropsychology, 34*(1), 31–42. https://doi.org/10.1037/neu0000576

Holtzer, R., Verghese, J., Wang, C., Hall, C. B., & Lipton, R. B. (2008). Within-person across-neuropsychological test variability and incident dementia. *JAMA, 300*(7), 823-830. doi:10.1001/jama.300.7.823

Holtzer, R., Wang, C. & Verghese, J. (2014). Performance variance on walking while talking tasks: theory, findings, and clinical implications. *AGE 36*, 373–381. https://doi.org/10.1007/s11357-013-9570-7

Hultsch, D. F., MacDonald, S. W., Hunter, M. A., Levy-Bencheton, J., & Strauss, E. (2000). Intraindividual variability in cognitive performance in older adults: Comparison of adults with mild dementia, adults with arthritis, and healthy adults. *Neuropsychology*, 14, 588 – 598. http://dx .doi.org/10.1037/0894-4105.14.4.588

Itabashi, R., Nishio, Y., Kataoka, Y., Yazawa, Y., Furui, E., Matsuda, M., & Mori, E. (2016). Damage to the Left Precentral Gyrus Is Associated With Apraxia of Speech in Acute Stroke. *Stroke*, *47*(1), 31–36. https://doi.org/10.1161/STROKEAHA.115.010402

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 : JINS*, *26*(6), 576–586. https://doi.org/10.1017/S135561772000003X

Koen, J. D., Srokova, S., & Rugg, M. D. (2020). Age-related neural dedifferentiation and cognition. *Current opinion in behavioral sciences*, *32*, 7–14. https://doi.org/10.1016/j.cobeha.2020.01.006

Leech, R. & Sharp, D. J. (2014). The role of the posterior cingulate cortex in cognition and disease. *Brain : a journal of neurology*, *137*(Pt 1), 12–32. https://doi.org/10.1093/brain/awt162

Lezak, M. D. (1995) Neuropsychological Assessment New York: Oxford University Press; 11.

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012) Neuropsychological Assessment. Oxford, New York: Oxford University Press

Li, Y., Li, P., Yang, Q. X., Eslinger, P. J., Sica, C. T., & Karunanayaka, P. (2017). Lexical-semantic search under different covert verbal fluency tasks: An fMRI study. *Frontiers in Behavioral Neuroscience. 11 (131*). doi: 10.3389/fnbeh.2017.00131

Lipton, R. B., Katz, M. J., Kuslansky, G., Sliwinski, M. J., Stewart, W. F., Verghese, J., ... Buschke, H. (2003). Screening for dementia by telephone using the memory impairment screen. *Journal of the American Geriatrics Society, 51,* 1382–1390. http://dx.doi.org/10.1046/j.1532- 5415.2003.51455.x

MacDonald, S. W., Nyberg, L., & Bäckman, L. (2006). Intra-individual variability in behavior: Links to brain structure, neurotransmission, and neuronal activity. *Trends in Neurosciences, 29,* 474-480. http://dx.doi.org/10.1016/j.tins.2006.06.011

Marqués-Iturria, I., Garolera, M., Pueyo, R., Segura, B., Hernan, I., García-García, I., … Jurado, M. Á. (2014). The interaction effect between BDNF val66met polymorphism and obesity on executive functions and frontal structure. *American journal of medical genetics. Part*

B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics, *165B*(3), 245–253. https://doi.org/10.1002/ajmg.b.32229

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*, 210-225. doi:10.1037/a0021531.

Monsch, A. U., Bondi, M. W., Butters, N., Paulsen, J. S., Salmon, D. P., Brugger, P., & Swenson, M. R. (1994). A comparison of category and letter fluency in Alzheimer's disease and Huntington's disease. *Neuropsychology, 8(1),* 25-30. 0894-4105/94/53.00

Moscovitch, M. (1994). Cognitive resources and dual task interference effects at retrieval in normal people: The role of the frontal lobes and medial temporal cortex. *Neuropsychology, 8*(4), 524-534. doi:10.1037/0894-4105.8.4.524

Murphy, K. J., Rich, J. B., & Troyer, A. K. (2006). Verbal fluency patterns in amnestic mild cognitive impairment are characteristic of Alzheimer's type dementia. *Journal of the International Neuropsychological Society : JINS*, *12*(4), 570–574. https://doi.org/10.1017/s1355617706060590

Nakamura-Palacios, E. M., Souza, R. S., Zago-Gomes, M. P., de Melo, A. M., Braga, F. S., Kubo, T. T., & Gasparetto, E. L. (2014). Gray matter volume in left rostral middle frontal and left cerebellar cortices predicts frontal executive performance in alcoholic subjects. *Alcoholism, clinical and experimental research*, *38*(4), 1126–1133. https://doi.org/10.1111/acer.12308

Pietrzykowski, M. O., Daigle, K. M., Waters, A. B., Swenson, L. P., & Gansler, D. A. (2022). The central executive network and executive function in healthy and persons with schizophrenia groups: a meta-analysis of structural and functional MRI. *Brain imaging and behavior*, *16*(3), 1451–1464. https://doi.org/10.1007/s11682-021-00589-3

Puoliväli, T., Palva, S., & Palva, J. M. (2020). Influence of multiple hypothesis testing on reproducibility in neuroimaging research: A simulation study and Python-based software. *Journal of neuroscience methods*, *337*, 108654. https://doi.org/10.1016/j.jneumeth.2020.108654

Raboutet, C., Sauzeon, H., Corsini, 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,* 268-280. doi:10.1080/13803390902984464.

Respino, M., Jaywant, A., Kuceyeski, A., Victoria, L. W., Hoptman, M. J., Scult, M. A., … Gunning, F. M. (2019). The impact of white matter hyperintensities on the structural connectome in late-life depression: Relationship to executive functions. *NeuroImage. Clinical*, *23*, 101852. https://doi.org/10.1016/j.nicl.2019.101852

Reuter, M., Schmansky, N. J., Rosas, H. D., & Fischl, B. (2012). Within-subject template estimation for unbiased longitudinal image analysis. *NeuroImage, 61(4),* 1402–1418. https://doi.org/10.1016/j.neuroimage.2012.02.084

Robinson, G., Shallice, T., Bozzali, M., & Cipolotti, L. (2012). The differing roles of the frontal cortex in fluency tests. *Brain : a journal of neurology*, *135*(Pt 7), 2202–2214. https://doi.org/10.1093/brain/aws142

Roye, S., Castagna, P. J., Calamia, M., De Vito, A. N., Lee, T. H., & Greening, S. G. (2020). Relationships between multiple dimensions of executive functioning and resting-state networks in adults. *Neuropsychologia*, *141*, 107418. https://doi.org/10.1016/j.neuropsychologia.2020.107418

Salat, D. H., Buckner, R. L., Snyder, A. Z., Greve, D. N., Desikan, R. S., Busa, E., … Fischl, B. (2004). Thinning of the cerebral cortex in aging. *Cerebral cortex (New York, N.Y.: 1991), 14(7)*, 721–730. https://doi.org/10.1093/cercor/bhh032

Schlösser, R., Hutchinson, M., Joseffer, S., Rusinek, H., Saarimaki, A., Stevenson, J., … Brodie, J.D. (1998). Functional magnetic resonance imaging of human brain activity in a verbal fluency task. *J. Neurol. Neurosurg. Psychiatry* 64, 492–498.

Ségonne, F., Dale, A. M., Busa, E., Glessner, M., Salat, D., Hahn, H. K., & Fischl, B. (2004). A hybrid approach to the skull stripping problem in MRI. *NeuroImage, 22(3),* 1060–1075. https://doi.org/10.1016/j.neuroimage.2004.03.032

Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans. Med. Imag. 17 (1),* 87-97. https://doi.org/10.1109/42.668698.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, *44*(1), 83–98. https://doi.org/10.1016/j.neuroimage.2008.03.061

Spreen, O., & Benton, A.L. (1977). Neurosensory center comprehensive examination for aphasia: *Manual of directions*. Victoria, British Columbia: University of Victoria

Stuss, D. T., Toth, J. P., Franchi, D., Alexander, M. P., Tipper, S., & Craik, F. I. (1999) Dissociation of attentional processes in patients with focal frontal and posterior lesions. *Neuropsychologia. 37*(*9*):1005–1027.

Stuss, D., Alexander, M., Hamer, L., Palumbo, C., Dempster, R., Binns, M., … Izukawa, D. (1998). The effects of focal anterior and posterior brain lesions on verbal fluency. *Journal of the International Neuropsychological Society 4*(03):265-278

Sung, K., Gordon, B., Vannorsdall, T. D., Ledoux, K., Pickett, E. J., Pearlson, G. D., & Schretlen, D. J. (2012). Semantic clustering of category fluency in schizophrenia examined with singular value decomposition. *Journal of The International Neuropsychological Society,* 18, 565-575. doi:10.1017/S1355617712000136.

Szatkowska, I., Szymańska, O., Bojarski, P., & Grabowska, A. (2007). Cognitive inhibition in patients with medial orbitofrontal damage. *Experimental brain research*, *181*(1), 109–115. https://doi.org/10.1007/s00221-007-0906-3

Tao, L., Zhu, M., & Cai, Q. (2020). Neural substrates of Chinese lexical production: The role of domain-general cognitive functions. *Neuropsychologia*, *138*, 107354. https://doi.org/10.1016/j.neuropsychologia.2020.107354

Troyer, A., Moscovitch, M., Winocur, G., Leach, L., & Freedman, M. (1998). Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease. *Journal of the International Neuropsychological Society, 4.* 137-143.

Wagner, S., Sebastian, A., Lieb, K., Tüscher, O., & Tadić, A. (2014). A coordinate-based ALE functional MRI meta-analysis of brain activation during verbal fluency tasks in healthy control subjects. *BMC neuroscience*, *15*, 19. https://doi.org/10.1186/1471-2202-15-19

Wang, Z., Yue, L., Cui, C., Liu, S., Wang, X., Li, Y., & Ma, L. (2019). Top-down control of the medial orbitofrontal cortex to nucleus accumbens core pathway in decisional impulsivity. *Brain structure & function*, *224*(7), 2437–2452. https://doi.org/10.1007/s00429-019-01913-w

Watanabe, E., Maki, A., Kawaguchi, F., Takashiro, K., Yamashita, Y., Koizumi, H., & Mayanagi, Y., (1998) Non-invasive assessment of language dominance with nearinfrared spectroscopic mapping. *Neurosci. Lett.256*(1), 49-52. https://doi.org/10.1016/S0304-3940(98)00754-X

Watson, P., van Wingen, G., & de Wit, S. (2018). Conflicted between Goal-Directed and Habitual Control, an fMRI Investigation. *eNeuro*, *5*(4), ENEURO.0240-18.2018. https://doi.org/10.1523/ENEURO.0240-18.2018

Weakley, A., Schmitter-Edgecombe, M., & Anderson, J. (2013). Analysis of verbal fluency ability in amnestic and non-amnestic mild cognitive impairment. Archives of Clinical Neuropsychology, 28(7), 721–731. doi:10.1093/arclin/act058

West, R., Murphy, K. J., Armilio, M. L., Craik, F. I. M., & Stuss, D. T. (2002). Lapses of Intention and Performance Variability Reveal Age-Related Increases in Fluctuations of Executive Control. *Brain and Cognition, 49*(3), 402-419. doi:10.1006/brcg.2001.1507

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

Wilkinson, G. S. (1993). *Wide Range Achievement Test: WRAT3* (rev., 3rd ed.). Wilmington, DE: Wide Range, Inc.

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *NeuroImage*, *92*(100), 381–397. https://doi.org/10.1016/j.neuroimage.2014.01.060

Zhang, S., Tsai, S. J., Hu, S., Xu, J., Chao, H. H., Calhoun, V. D., & Li, C. S. (2015). Independent component analysis of functional networks for response inhibition: Inter-subject variation in stop signal reaction time. *Human brain mapping*, *36*(9), 3289–3302. https://doi.org/10.1002/hbm.22819

Yuan, P., & Raz, N. (2014). Prefrontal cortex and executive functions in healthy adults: a meta-analysis of structural neuroimaging studies. *Neuroscience and biobehavioral reviews*, *42*, 180–192. https://doi.org/10.1016/j.neubiorev.2014.02.005

Table 1. Descriptives of target variables.

		Mean (SD)	Range
Age years		74.91 (5.31)	65 – 89
Education years		15.5 (3.22)	7 – 20
GHS		1.30 (1.04)	0 – 4
WRAT3	Raw	35.27 (5.90)	17 – 42
	Standard Score	108.94 (9.66)	79 - 120
Total Letter Fluency	Raw	13.97 (6.70)	0 – 33
	Z-score	0.23 (1.10)	-2.87 – 2.97
Total Category Fluency	Raw	14.48 (8.25)	2 – 39
	Z-score	0.30 (1.46)	-2.74 – 4.43

Note: GHS = Global Health Score; WRAT3 = Wide-Range Achievement Test.

Table 2. Verbal Fluency word generation across intervals.

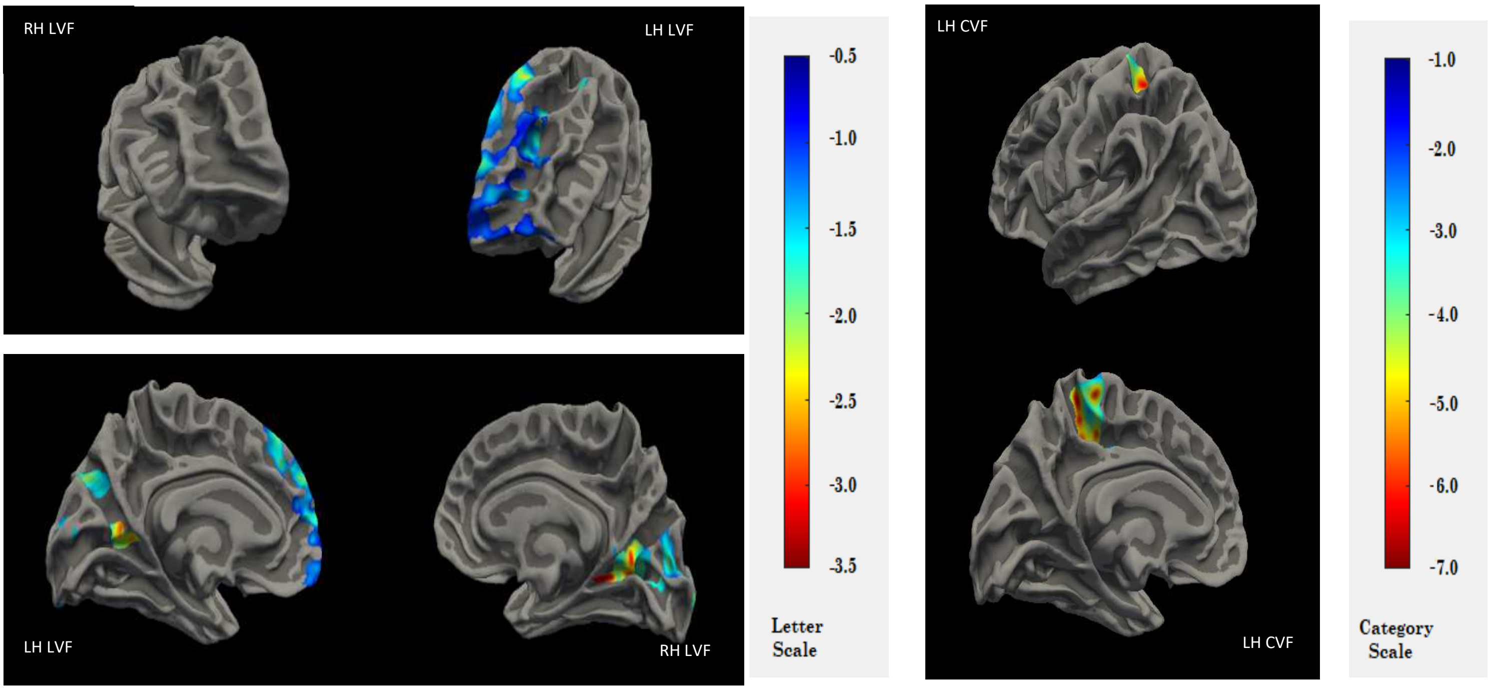
		Mean (SD)	Range
Letter Fluency	0-20-sec	6.60 (2.01)	1.67 – 11.00
	21-40-sec	4.10 (1.68)	0.33 – 8.67
	41-60-sec	3.27 (1.45)	0.00 – 7.33
Category Fluency	0-20-sec	7.81 (8.00)	3.00 – 13.00
	21-40-sec	3.88 (1.49)	1.00 – 7.33
	41-60-sec	2.78 (1.47)	0.671 – 8.00

Table 3. Significant areas of GMV implicated in change in word generation across time intervals in letter and category VF.

		Left Hemisphere			Right Hemisphere		
	Region	Standardized Beta (words/cc)	*p*	Volume (cc)	Standardized Beta (words/cc)	*p*	Volume (cc)
		Letter Fluency					
Frontal	Caudal middle frontal	-1.782	0.047	291.05			
	Frontal pole	-0.725	0.042	838.41			
	Lateral orbitofrontal	-0.723	0.046	63.07			
	Medial orbitofrontal	-0.846	0.044	767.44			
	Pars orbitalis	-0.856	0.046	194.70			
	Rostral middle frontal	-0.868	0.044	4398.18			
	Superior frontal	-1.241	0.042	5539.69			
Parietal	Isthmus cingulate	-2.384	0.045	148.22	-3.046	0.038	65.15
	Precuneus	-2.043	0.045	848.48	-2.060	0.031	2.99
	Superior parietal	-1.714	0.048	162.95			
Occipital	Cuneus	-1.531	0.049	227.59	-1.486	0.032	812.37
	Lateral occipital	-1.541	0.049	333.77	-1.474	0.048	147.84
	Lingual	-2.178	0.048	11.60	-2.539	0.032	110.08
	Pericalcarine	-1.659	0.049	36.75	-1.669	0.038	394.12
		Category Fluency					
Frontal	Paracentral	-4.985	0.038	1623.69			
	Precentral	-2.203	0.045	72.65			
Parietal	Postcentral	-4.485	0.041	362.97			
	Posterior cingulate	-5.160	0.041	401.47			
	Precuneus	-8.405	0.045	5.67			
	Superior parietal	-6.438	0.045	26.12			

Note: Significant main effects of volume and time described in text.

Figure 1. Lateral and medial views of unstandardized beta values in significant areas of the cortex where reduced GMV is associated with attenuated word generation slope as identified by the linear mixed effects models, projected onto the “fsaverage” cortical surface from the Freesurfer package.



NB: RH LVF (right hemisphere letter verbal fluency); LH LVF (left hemisphere letter verbal fluency); LH CVF (left hemisphere category verbal fluency)

(Include color)