Psychology Research and Behavior Management

Dovepress

open access to scientific and medical research

Open Access Full Text Article

REVIEW

Current Understanding of Verbal Fluency in Alzheimer's Disease: Evidence to Date

Laura M Wright[1], Matteo De Marco[2], Annalena Venneri[2,3]

[1]Translational and Clinical Research Institute, Newcastle University, Newcastle Upon Tyne, UK; [2]Department of Life Sciences, Brunel University London, London, UK; [3]Department of Medicine and Surgery, University of Parma, Parma, Italy

Correspondence: Annalena Venneri, Email annalena.venneri@brunel.ac.uk

Abstract: Since their development, verbal fluency tests (VFTs) have been used extensively throughout research and in clinical settings to assess a variety of cognitive functions in diverse populations. In Alzheimer's disease (AD), these tasks have proven particularly valuable in identifying the earliest forms of cognitive decline in semantic processing and have been shown to relate specifically to brain regions associated with the initial stages of pathological change. In recent years, researchers have developed more nuanced techniques to evaluate verbal fluency performance, extracting a wide range of cognitive metrics from these simple neuropsychological tests. Such novel techniques allow for a more detailed exploration of the cognitive processes underlying successful task performance beyond the raw test score. The versatility of VFTs and the richness of data they may provide, in light of their low cost and speed of administration, therefore, highlight their potential value both in future research as outcome measures for clinical trials and in a clinical setting as a screening measure for early detection of neurodegenerative diseases.

Keywords: fluency, Alzheimer's disease, AD, semantic, mild cognitive impairment, MCI

Since their development in the late 1960s, verbal fluency tests (VFTs) have become a widely used measure in both cognitive psychology and neuropsychological disciplines.[1,2] Being both cost effective and quick to administer, VFTs represent accessible bedside screening measures for the assessment of cognitive decline. As such, they are now a staple of neuropsychological test batteries, used both in clinical settings and research. Specifically, verbal fluency refers to a task in which participants are required to list as many words adhering to a given criterion as possible. Most commonly, this includes listing words beginning with a certain letter (e.g., F), as in phonemic, or letter fluency,[1] or belonging to a given category (e.g., animals), as in semantic, or category fluency,[2] in a specific time constraint. During both the letter fluency task (LFT) and the category fluency task (CFT), controlled retrieval processes, thought to be mediated by frontal lobe structures,[3,4] are vital for successful performance. As such, VFTs are often included as measures of executive functioning when assessing neuropsychological function in Alzheimer's disease (AD) and other cognitively impaired groups. Unlike LFTs however, the CFT additionally relies upon access to one's vocabulary and previously learned semantic associations represented within a semantic memory store thought to be sustained by a neural network converging within the temporal lobes.[4–7] Concurrent use of multiple VFTs, such as CFTs and LFTs, is, therefore, especially useful for isolating the multiple cognitive processes involved in each to ascertain the presence of a controlled retrieval deficit vs. semantic access impairment.[8]

Patterns of Verbal Fluency Performance in Alzheimer's Disease: Category and Letter Fluency

As the most common cause of dementia in older adults, considerable research has aimed to understand the impact of AD related cortical degradation on measures of verbal fluency. It is now well documented that, although both types of verbal fluency are susceptible to AD type neurodegeneration, even in the earliest stages[9], the CFT appears to be consistently impaired to a greater extent, suggesting that degradation of the semantic system in this cohort significantly outweighs

Received: 31 December 2022
Accepted: 5 April 2023
Published: 5 May 2023

disruption of executive functions[10]. One of the first studies to outline such a discrepancy in AD patients was carried out by Monsch et al[11] who found that the CFT was best able to differentiate AD patients from healthy aging controls, with a sensitivity of 100% and specificity of 92.5%, in contrast with the LFT that only reached a sensitivity and specificity of 89% and 85%, respectively. Since this seminal paper, several studies have recorded a similar discrepancy between the relative deficits in each of these forms of VFT in AD patients. In a meta-analysis conducted by Henry, Crawford, and Phillips[10], it was concluded, when considering studies using VFTs, alongside several other neuropsychological measures, that AD patients present with significantly more impairment on CFTs, as well as other tasks of semantic memory, when compared with LFTs. Furthermore, semantic fluency deficits and not phonemic fluency deficits, were considered a differential deficit, independent of declines in verbal intelligence or psychomotor speed. Taken together, these findings suggest that the CFT deficit present in AD is likely to reflect considerable damage to the semantic memory store rather than a disruption of overall controlled retrieval. Finally, the authors report that semantic fluency was generally impaired to a greater extent than naming, suggesting that the added demands of effortful retrieval in this task make it a more sensitive marker of semantic memory decline than a simple naming task, and this observation has significant implications for its use in clinical assessment early in the course of disease.

Further evidence from research in patients with mild cognitive impairment (MCI) has since described a similar pattern of VFT impairment in this prodromal population. Murphy et al[12] administered LFT and CFT trials to three participant groups consisting of individuals with amnestic MCI (aMCI), considered a prodromal stage of AD,[13] AD dementia and healthy controls. Using a cross-sectional design, this study was able to demonstrate clear differences in VFT performance patterns across diagnostic stages. In accordance with normative data,[14] healthy older adults demonstrated a significant advantage in the semantic condition relative to the phonemic condition. However, in the patient groups, this advantage disappeared, with MCI patients showing only a marginal, non-significant, semantic advantage and dementia patients showing the reverse pattern, with a significant phonemic advantage. More recent evidence from Chasles et al,[15] has similarly demonstrated this pattern in one-to-one matched groups, again including aMCI patients, dementia patients and healthy controls. Again, a significant semantic advantage was evidenced in the control group that fell below significance in the aMCI group and was virtually non-existent in dementia patients. Furthermore, in the LFT, aMCI patients were found to perform similarly to controls despite showing significant impairments in the CFT. By separating MCI groups into amnestic and non-amnestic, and further categorizing those as single or multiple domain, a study by Liampas et al[16] was able to assess patterns of VFT performance as they manifest depending on the type (amnestic vs. non-amnestic) and severity of cognitive impairment (i.e., from single to multiple domain). In line with previous findings, aMCI patients tended to perform significantly worse than those with a non-amnestic profile (naMCI) on the CFT, despite performing similarly on the LFT, suggesting a differential deficit of semantic memory function in aMCI, perhaps relating to the increased likelihood of such individuals to develop AD type dementia.[13] Furthermore, when categorized according to a single or multiple domain profile, this differential deficit was only appreciable between the single domain aMCI and naMCI groups. In patients with a moderate cognitive impairment, reflected in multiple domains, aMCI and naMCI patients did not differ on either VFT. Such findings suggest that patients with an amnestic profile demonstrate greater semantic deficits than those with naMCI even at a minimal level of impairment, reflected by a single domain cognitive deficit. In patients with a multiple domain cognitive profile however, such differential deficits are lost possibly owing to similar levels of impairment in extraneous cognitive domains such as executive function and processing speed that are related to performance in both VFTs.

Using longitudinal data, Clark et al[17] were able to document this decline in semantic fluency as it occurs throughout disease development. They found that both preclinical patients who developed dementia during follow up and patients who presented with dementia at baseline showed a significantly greater rate of decline in semantic fluency over time, relative to phonemic fluency. Furthermore, at baseline both groups were significantly more impaired in the CFT relative to the LFT when compared with a group of cognitively normal controls who remained healthy throughout follow up, with the preclinical group performing comparatively with cognitively normal individuals on the LFT. Although in this study greater semantic decline was also evident among the cognitively normal group, despite the relative preservation of semantic memory in aging,[14,18] this effect was significantly exacerbated in the preclinical AD group. In the AD group both VFTs showed significant decline, but the degradation of semantic fluency continued to decline at a much faster rate.

These findings indicate that, although declines in CFTs may be present in normal aging, this is significantly accelerated by AD pathology and such acceleration may be observable several years prior to diagnosis. Clark et al further comment that, given the known presence of AD pathology even decades prior to the dementia stages,[19] there is a possibility that members of the control group included in this study may have already been experiencing the initial stages of AD pathology despite appearing cognitively normal on neuropsychological testing. As semantic memory function is thought to be relatively well preserved in healthy aging populations,[14,18] this may indicate that semantic fluency could be a sensitive measure to the very earliest stages of pathological decline.

Longitudinal studies have indicated strong prognostic value for CFTs in predicting cognitive decline and dementia development in MCI and even cognitively healthy individuals. Recent work by García-Herranz et al[20] found that, cross-sectionally, CFTs performed very well both in discriminating between healthy controls and MCI patients, as well as between MCI converters and non-converters. Although all the VFTs included performed well in discriminating the groups in this study, the LFT was the least effective, particularly when discriminating between MCI patients who later progressed to dementia and those who did not, with a sensitivity and specificity of 57.8 and 66.7 respectively (values expressing percentages), compared with the much higher accuracy of the animal fluency task (sensitivity = 75.6 and specificity = 81.5). When entering multiple VFTs into stepwise logistic regression models, the animals test alone performed best, with no other fluency measure, including both phonemic fluency and other category tasks, having significant power to differentiate between groups. Similarly, in their 2018 study, Gallucci et al[21] found that poor performance on two measures of verbal memory, namely the Rey Auditory Verbal Learning Task and the CFT, was associated with a greater risk of decline to dementia in aMCI patients, with odds ratios of 0.58 and 0.81 respectively. As in García-Herranz et al,[20] the predictive value of LFTs, as well as performance in other cognitive domains, however, was found not to relate significantly to dementia risk. Population-based studies have further demonstrated that CFTs have strong prognostic value for predicting cognitive decline to both aMCI and dementia in cognitively normal individuals, even in middle-age,[22–24] a finding that again was not reflected in phonemic fluency measures in any study. A longitudinal study, conducted by Vonk et al[25] similarly found lower baseline performances and faster rates of decline in a CFT among non-demented individuals who presented with incident dementia or aMCI at follow-up, when compared with those who had no cognitive impairment or a non-amnestic form of MCI. Furthermore, similar differences were seen at baseline, and in the rate of semantic decline, among ApoE-ε4 carriers compared with non-carriers and those with a score on the Clinical Dementia Rating of 0.5, compared with a Clinical Dementia Rating of 0. Notably, despite evident declines in semantic fluency among groups at-risk for AD (i.e., ApoE-ε4 carries, participants with a Clinical Dementia Rating of 0.5, aMCI and incident dementia groups), only the group evidencing dementia at follow-up demonstrated declines in the LFT that, in line with previous research, was to a lesser extent than CFT declines. Neurodegenerative measures relating to AD, including lower hippocampal volumes, increased white matter hyperintensities and overall cortical thinning, as well as reduced metabolic functioning within a number of AD related areas, including the entorhinal cortex, inferior parietal lobule and posterior cingulate gyrus/precuneus, were all correlated with lower CFT performance at baseline and similar findings were seen in relation to rates of CFT performance decline. Conversely, no such relationships were found between baseline LFT scores and cortical signatures of AD and, although global indices of neurodegeneration were correlated with the rate of LFT decline, this showed no specificity for AD-type alterations. Taken together, these findings therefore suggest, as per the conclusion of previous studies, that declines in semantic fluency, particularly in the absence of phonemic fluency decline, represent a distinct marker for AD degeneration early in the course of disease that may predict subsequent development of aMCI or AD dementia.

In line with Vonk et al,[25] further evidence from studies utilizing at-risk groups, such as carriers of the ApoE-ε4 allele, and cognitive unimpaired individuals testing positive for Amyloid Beta (Aβ) accumulation, has demonstrated changes in CFT performance that are detectable even in cognitively healthy individuals who do not yet show any additional signs of neurodegeneration.[26–28] A study from Papp et al[27] found that clinically normal older adults who tested positive for the deposition of Aβ (Aβ+) declined to a significantly greater extent longitudinally on a CFT than Aβ negative (Aβ-) participants. Although Aβ+ participants also showed similarly greater declines in LFT performance relative to the Aβ- group, this difference did not retain significance when CFT performance was added as a covariate. Greater declines in semantic fluency among the Aβ+ group retained significance even when covarying for phonemic fluency, however,

suggesting that additional declines in semantic retrieval processes unique to CFTs could not be accounted for by disruptions to executive control. Furthermore, Papp et al went on to suggest, in a later study, that CFTs could identify unique variances in cognitive decline measured in relation to Aβ accumulation, with Aβ+ individuals demonstrating continued significant decline relative to Aβ- in CFTs even when controlling for the overall decline measured by the preclinical Alzheimer's cognitive composite (PACC).[28] Moreover, the removal of CFTs from the PACC in this study resulted in a 20% reduction in the measured longitudinal amyloid-related decline at 3 years of follow-up, making semantic fluency one of the greatest contributors to variance in PACC measured cognitive function found in this study. Further research looking at the relationship between VFTs and total tau levels has also demonstrated significant correlations between this cerebrospinal fluid biomarker and semantic fluency performance among MCI patients,[29] suggesting that this measure is indicative of both preclinical and prodromal accumulations of AD related protein aggregates. As one of the first cognitive changes detectable up to 12 years prior to a dementia diagnosis,[30] evidence establishing a link between pathological protein accumulation and CFT performance is, therefore, further indicative of the value of this task for identifying prodromal and even preclinical manifestations of AD.

Despite extensive evidence supporting the use of CFTs in AD diagnosis, some issues remain. Semantic knowledge is inherently related to the idiosyncratic experience of the individual. As such, tests examining different categories (i.e., animals, vegetables, clothing, etc.) may elicit both qualitatively and quantitatively different responses depending on both language and cultural background.[31–34] Even among homogenous participant groups, questions remain regarding category-specific deficits in AD. In studies of semantic memory in AD, categories of living things such as animals, fruits or vegetables have frequently been shown to be impaired to a greater extent and have a higher level of diagnostic accuracy,[35–38] although such findings have not been corroborated in all cases.[39] A recent study has further suggested that even within the living things categories, the sub-category of vegetables may be more strongly associated with underlying brain network function and dementia severity than the commonly used animal fluency task.[40] Use of CFTs in clinical practice. therefore. requires further research not only identifying the best categories to use for detecting early AD, but further to develop culturally specific normative data so that these tests may be applied worldwide.

Alternative Measures of Verbal Fluency Performance

Item-Based Scoring

While the CFT was historically designed as a measure of free recall,[41] its performance is also influenced by functions other than lexical/semantic processing and semantic memory. These include executive functioning,[42] processing speed,[43] speed of lexical access[44] and episodic memory.[45] Accordingly, factors analyses have found CFT performance to be explained by each of these factors to a greater or lesser extent.[42,46–48] As a result, the standard CFT score must be considered, inevitably, as a multidimensional index. Although this has little or no effect on test validity at a group level, it is difficult to evaluate whether two participants (or patients) who name an equal number of exemplars have relied on the same cognitive resources. Gender differences, for instance, exist in the ability to rely on a switching-based strategy.[49] Anecdotal evidence, moreover, indicates that a small proportion of participants decides to explore large categories such as animals via a letter-by-letter strategy, i.e., by naming words like "aardvark", "armadillo" or "antelope" at the start of the performance, moving on, then, to entries starting with B, and progressing in alphabetical order at least for the first half of the minute[50] (Supplementary Material).

The inability to pinpoint the contribution of semantic memory to CFT performance in the single individual has important clinical implications. As semantic memory alterations appear to be the earliest significant changes in preclinical AD,[51] it is of primary importance to identify methodologies that are capable of quantifying these subtle changes from CFT performance with more precision. It is with this objective in mind that "item-based" approaches have been proposed. Item-based methods do not look at the overall count of the correct entries, but evaluate the qualitative properties of individual words. This is based on the principle that a better-preserved semantic memory will allow the individual to retrieve "more difficult" exemplars of that category. In line with this premise, researchers have explored various data-driven methods to quantify "how difficult" each word is within its own category. A simple approach, in this respect, is to assess general word frequency *within the sample itself*, i.e., by rating each word as a function of the

proportion of participants within the sample that generated it as a measure of originality.[52] Despite its parsimony, this method to score item-based difficulty is *relative* (or sample-dependent) and, arguably, it does not allow direct comparisons of performance between groups acquired in different experimental or clinical contexts (e.g., neurological patients and undergraduate students). Conversely, a number of initiatives have assessed item-based difficulty on large samples, to obtain *absolute* (or sample-independent) ratings that can be consistently used as normative data across studies. Some of these were calculated by retrospective analyses of CFT word lists, as done, for instance, by Quaranta et al[53] (to quantify category-specific "typicality" of words) or bodies of published (or publicly available) text-based material such as television subtitles, social media, books or internet blogs, to generate measures representative of words' "frequency of use".[54] Other aspects of word difficulty have instead been defined and computed prospectively (rather than retrospectively), by asking participants recruited *ad hoc* to judge properties of individual words by using a Likert scale as carried out, for instance, for affective ratings of "valence", "arousal" and "dominance",[55] or to report an autobiographical detail associated with each word, as done in relation to words' "age of acquisition".[56,57]

Although the above studies are based on a range of diverse methodologies (and most were designed for the purpose of providing normative data that could be used in a range of tests and tasks, not exclusively for the CFT), researchers interested in the study of normal and abnormal aging have exploited item-based ratings to characterize the profile of CFT performance either by scoring all generated words (and calculating an overall average), or by focusing on the first segment (i.e., first 5 words) of performance.[56,58] The purpose of the latter method is to assess the aspects of CFT performance that are as little affected by executive control mechanisms as possible, as the earliest words generated are typically the result of automatic memory retrieval.[59] This allows a fair comparison of patients and controls, for instance, since patients may suffer from a concurrent deficit in executive processing that may prevent them from deploying the adequate resources that are needed to explore a category more in depth.

Clinical studies found that item-based scoring methods are sensitive to diagnostic status. For instance, age of acquisition of CFT words is significantly lower in patients with a clinical diagnosis of AD when these are compared with healthy controls[56,60] or with patients with a functional, non-neurodegenerative etiology.[58] Comparably, other studies have found that patients with AD tend to generate words that are more typical of their category than healthy controls, regardless of whether typicality scores are generated with sample-dependent[61] or sample-independent[62] methodologies. Importantly, not all item-based indices of difficulty appear to differ between patients and controls as, for instance, no effect was reported for words' frequency of use in some of these studies.[56,62] A study carried out in a cohort of non-demented adults, however, reported that frequency of use differs between carriers and non-carriers of the ApoE-ε4 risk factor,[63] with carriers generating, on average, words that are significantly more frequent. This suggests that item-based difficulty is a multifaceted construct, and that a profile of features (rather than a single feature) could help characterize semantic alterations in more detail, not only as a function of the diagnosis of interest, but also as a function of a series of variables that contribute to the overall clinical phenotype. More studies are required to investigate the mechanistic link between AD and item-based scoring of CFT words. Typicality of CFT words has been found associated with gray-matter volumes in the perirhinal cortex,[64] a finding that suggests a specific link with the preclinical phase of AD, but more evidence is needed to substantiate the link between individual item-based features and neuroanatomical integrity.

Item-based approaches have not been limited to the scoring of individual words but have also informed the study of more intricate properties of retrieval. Quaranta et al,[65] for instance, investigated the "semantic relatedness" between words generated consecutively during the CFT, on the assumption that individuals with underlying subtle difficulties in semantic-memory processing would show alterations in the way in which categories are explored. In measuring "path length" (i.e., edge-based distance between two words as informed by the normative network analysis of the lexicon carried out by the WordNet initiative – https://wordnet.princeton.edu/) and "extended gloss overlap" (i.e., the number of terms shared by the definition of two words), the researchers characterized the profile of two CFT categories, reporting significant differences between controls and patients with MCI who would later progress to dementia.[65]

A further approach to item-based scoring is that focusing on how item-based ratings change during the course of the CFT minute. Since more difficult entries tend to be generated as the task progresses,[52,66] it is of interest to operationalize and quantify the tendency shown by participants/patients to explore the target category in more detail. To this end, De Marco et al[67] proposed a scoring method focused on the statistical correlation between item-based ratings (semantic and

non-semantic features were explored) and the order of word retrieval, i.e., an ordinal variable ranging from 1 to *n* indicating the serial position of the word within the list of retrieved entries. They documented an age-dependent effect on words' valence (i.e., indicating how pleasant each word is), with both older and younger adults generating words that are increasingly less pleasant, but with younger adults showing a trend characterized by a significantly steeper slope.[67] In a follow-up study, in addition, this team of authors investigated the neural substrate of this scoring method, reporting significant associations between four of these correlational indices (i.e., those involving typicality, frequency of use, dominance and sensorimotor interaction) and aspects of the default-mode network that are linked to semantic control.[50] The idea that semantic memory and other cognitive functions do not sustain task performance in a regular and constant way during the course of the whole testing minute is reflected by other procedural approaches to CFT scoring. Other than the aforementioned approach based on the characterization of the first 5 retrieved words,[56,58] other authors[68] have proposed splitting the global count into four sub-scores, each related to a 15-sec interval. This is to characterize task performance in a way that is more heterogeneous (as there is evidence that the earliest words are more dependent on automatic, rather than controlled semantic memory processing),[59] yet reliant on a scoring procedure that is of easy implementation.

In summary, item-based scoring methods are promising to quantify aspects of semantic memory retrieval that may be obfuscated when the sole standard score based on word count is considered. A large amount of experimental work, however, still needs to be carried out to describe the potential of this approach in more detail.

The Semantic vs. Phonemic Fluency Impairment Discrepancy

In recent years, the use of discrepancy scores, measuring the degree to which participants' performances differ on CFTs and LFTs, has, in a less nuanced approach than item-based methods, been used as a simple means to isolate the semantic component of VFTs and identify an AD-like cognitive impairment. A study by Marra et al[69] operationalized this discrepancy as the semantic-phonological delta (SPD), a measure that calculates the differences between the raw fluency scores in each domain (CFT–LFT) while accounting for the number of tasks (letters and categories), where a negative SPD reflects lower performance in the CFT than the LFT. Their results showed that in aMCI, as well as mild and moderate AD dementia groups, the SPD was significantly lower than in controls, suggesting, as expected, greater impairments in semantic fluency. Moreover, this measure of verbal fluency discrepancy was found to be significantly predictive of conversion to dementia in a sub-sample of MCI patients followed up over five years, with those with an SPD below the median score for this group demonstrating markedly shorter times to conversion than those with scores above. Evidence from Vaughan et al,[70] has similarly demonstrated that discrepancies in verbal fluency performance, reflecting a reduction in the semantic advantage, may provide a potential predictive indicator of progression to dementia among MCI patients. In this case, results showed that with each unit decrease in discrepancy scores, the odds of progressing to dementia increased, a finding that was not reflected for the individual CFT and LFT scores. In contrast, however, an earlier cross-sectional analysis by Teng et al,[71] using the same SPD calculation as later outlined by Marra et al,[69] found that differentiation of controls from aMCI and AD dementia patients was most accurate when using the CFT scores alone (Area under the curve [AUC] distinguishing controls from dementia = 0.96) and least accurate using discrepancy scores (AUC = 0.73). In accordance with both previous and subsequent research, however, the results of this study similarly found that aMCI and AD dementia patients demonstrated lower discrepancy scores not only when compared with healthy controls but further in comparison with naMCI patients.

Despite semantic-phonemic fluency discrepancies now representing a well-described phenomenon in AD, there remains some conflict within the literature. Cross-sectional research into the diagnostic utility of such discrepancies, using sensitivity and specificity calculations, has not demonstrated the same predictive power suggested by longitudinal analyses[71,72] and some studies in prodromal groups have failed to identify this pattern at all in MCI stages.[73,74] Similarly, a meta-analysis by Laws et al concluded that, in the fifty studies chosen for their analysis, effect sizes did not differ between the 2167 controls and 1771 AD patients included when determining performance differences between CFTs and LFTs.[75] Such conflict may arise from differences in study design, such as the specific fluency criteria chosen (i.e., which letters or categories are used), that has been found to have a profound effect on patterns of performance in healthy aging and AD,[76] or from differences in how MCI is operationalized in each study, particularly if participants were categorized

according to a single-domain amnestic deficit. One factor, however, that may further contribute to such conflict is a simple limitation in the methodology of many studies exploring this phenomenon in AD: the use of the raw number of correct words produced on each task to calculate discrepancy scores. The limitation of using raw fluency scores to evaluate performance discrepancies is that this method does not provide information as to how far patients have declined on each test relative to their normal function or to the normal expected function for individuals of their age and education level. To illustrate this, a raw score of 12 on a LFT and 11 on a CFT may seem, at face value, a relatively insignificant difference in performance. However, in healthy adults, VFTs are often characterized by a significant semantic advantage.[14] If normative data for this age and education level, therefore, produced a mean of 22 on a LFT and 32 on a CFT, then the amount of decline relative to controls would be far greater in semantic fluency than phonemic, providing a more clinically meaningful measure of semantic memory decline in cognitive impairment. This could, therefore, explain why the AD patients in Cerhan et al's study[72] were found to differ significantly from controls in terms of discrepancies, but that this discrepancy was not useful for predicting group membership. This would also explain why, across the 135 studies included, Laws et al[75] found significantly greater effect sizes when assessing the degree of impairment on CFTs in AD, compared with LFTs, despite showing no difference in effect sizes for fluency discrepancies between patients and controls in the smaller fifty study sample. In prodromal groups, studies utilizing normative data from healthy controls to determine discrepancy scores based on standardized fluency *z* scores have not only demonstrated the typical disproportionate semantic fluency impairment expected in AD,[71,77] but have further demonstrated that such discrepancies can distinguish healthy controls from MCI patients with a relatively high sensitivity and specificity, that, in a study by Lonie et al,[78] was found to be marginally higher than a word list learning task typically used in clinics. Furthermore, evidence from longitudinal studies looking at rates of decline, rather than static differences in raw fluency scores, have consistently shown significantly greater rates of decline in CFTs compared with LFTs,[17,25,70,79] suggesting that disproportionate performances on VFTs are most accurately identified when reflecting the relative distance from baseline functioning or when operationalized as the relative difference between deviations from an established norm.

Clustering and Switching

Two behavioral components that are well established as essential to successful fluency performance are clustering and switching. First operationalized by Troyer, Moscovitch and Winocur in 1997,[80] clustering refers to the generation of words within a sub-category, while switching describes one's ability to switch to new sub-categories. Thought to reflect effortful executive control processes, switching has been found to relate primarily to frontal lobe structures, while clustering, which is conversely a reflection of automatic semantic association, is thought to be mediated by structures within the temporal lobe.[81,82] The dissociation of these two behavioral strategies and the heavy involvement of temporal lobe structures in AD has led researchers to explore the components of clustering and switching as they manifest across the AD spectrum and evaluate their potential as a diagnostic indicator for AD related cognitive decline. In their early study, Troyer et al[82] found a reduction in mean cluster size and number of switches, on the CFT only, in patients with AD dementia when compared with healthy older adults. Since then, several studies have suggested that changes to clustering and switching processes, in CFTs in particular, are present in AD, even in prodromal stages, although the nature of such changes remains unclear, with authors arguing both for a combined deficit of clustering and switching strategies in AD,[83,84] and others presenting evidence to support the existence of a singular primary deficit in either clustering[85] or switching[86] at a given disease stage. Despite the conflict within the literature, evidence suggests that impairments in these performance strategies may occur many years prior to dementia diagnosis[85,86] and may, therefore, potentially contribute to earlier diagnosis and prediction of incident AD dementia.

Fluency Errors

Total scores on VFTs are usually calculated as the number of unique words produced correctly according to their given letter or category. The two main error types on fluency tasks are, therefore, classified into perseverations, in which an individual may repeat a previously given response, and intrusions, in which a word is given that does not adhere to the given criterion (i.e., a word that begins with an incorrect letter or is not an exemplar of a given category). Perseverations in particular have been found to be significantly more common in AD than in healthy aging, even in the mildest stages of

dementia.[87] Errors on VFTs have, therefore, been investigated as a means to determine qualitative changes in fluency performance that may be appreciable in the earliest disease stages prior to the onset of dementia and a significant change on traditional VFT scoring measures. In a study by Pakhomov et al,[88] perseverations on the animals CFT was found to relate significantly to incident cognitive impairment in a large cohort of cognitively normal individuals. A similar study by Liampas et al,[89] utilizing data from the population-based Hellenic Longitudinal Investigation of Aging and Diet cohort, corroborated these findings, showing that the number of perseverations, but not intrusions, present on a CFT was strongly related to incident all-cause and AD dementia at follow-up among cognitively normal adults. Such findings, therefore, suggest that although intrusion errors may have little value in the detection of AD type cognitive decline, a simple count of perseveration errors may provide an additional screening measure for early detection of AD related change in cognitive function.

Neural Underpinnings of Verbal Fluency Performance in Alzheimer's Disease

Semantic Fluency

Given the interest surrounding VFT deficits as part of the AD neuropsychological phenotype, there has now been a substantial body of work aiming to uncover the neural correlates of such changes. Semantic fluency deficits in AD have been found to relate to both structural and functional changes in brain areas including medial and lateral temporal lobes,[64,90,91] frontal regions such as the anterior cingulate, prefrontal cortex, and superior and middle frontal gyri,[92,93] as well as inferior parietal lobules.[94] Similar regions, largely involving temporal lobes, but also extending into frontal areas, have since been implicated in MCI patients,[95,96] with CFT impairments in this population correlating with gray matter integrity in inferior and medial temporal regions,[97,98] even when patients present with no significant decreases in gray or white matter volumes relative to controls.[99] Moreover, performance on CFTs has, both in dementia and prodromal groups, been related specifically to gray and white matter integrity within perirhinal and entorhinal cortices known to be affected by the very earliest manifestations of AD pathology (see Table 1).[64,98–101] Only very few studies into CFT neural correlates in AD have analyzed the relationship between fluency and disease-related changes in white matter, with the majority focusing on gray matter volumes and functional connectivity. In their 2016 study however, Rodríguez-Aranda et al[102] found that, although accuracy on both VFTs correlated with white matter integrity, measured by fractional anisotropy (FA) and mean diffusivity (MD), in widespread areas including corpus callosum, forceps minor, left inferior fronto-occipital fasciculus and left superior longitudinal fasciculus, CFT scores were uniquely associated with FA values in the right inferior fronto-occipital fasciculus and right superior longitudinal fasciculus, and MD values in the posterior

Table 1 Neuropathological Associations with VFTs Across the Clinical Stages of AD

Clinical Stage	Task	Areas Involved	Evidence
MCI	CFT	• Perirhinal cortex • Entorhinal cortex • Hippocampus • Inferior temporal lobe	Loewenstein et al, 2017[97] Venneri et al, 2019[98] Wright et al, 2022[77]
Moderate MCI/ Mild AD Dementia	CFT	• Parahippocampal gyrus • Wider medial and lateral temporal lobes	Venneri et al, 2008[64] Wright et al, 2022[77]
Moderate AD Dementia	CFT and LFT	• Extensive temporal involvement • Posterior temporal lobes • Temporo-occipital cortex • Parietal association cortices • Parieto-occipital cortex • Frontal regions including premotor and dorsolateral prefrontal cortices as well as lateral and medial frontal cortices	Apostolova et al, 2008[92] Dos Santos et al, 2011[93] Eastman et al, 2013[94] Rodríguez-Aranda et al, 2016[102] Wright et al, 2022[77]

part of the right cingulum. Such findings are supportive of the notion that VFTs in AD are associated with white matter integrity within extensive areas involved in the left lateralized language network[103] and that declines in processing speed may contribute to cognitive impairment in this task.[43] In CFTs specifically, however, task deficits in AD were also related to white matter within the right hemisphere, suggesting that semantic fluency decline in these patients is likely a further reflection of the degradation of right-sided structures thought to be more heavily relied upon in aging to support semantic functioning.[104]

Phonemic Fluency

It is generally understood that although both fluency measures elicit multiple executive functions involved in controlled retrieval including appropriate word selection, updating and monitoring of working memory and inhibition of incorrect responses, LFTs, which involve additional processing due to the need to suppress automatic semantic associations and identify words based on an orthographic cue, are thought to place slighter greater demands on executive control systems.[44] In line with this, lesion studies have cited frontal lobe structures as being integral to both fluency processes, but LFTs in particular, and temporal structures as mediating semantic fluency performance.[4,105,106] Similarly, neuroimaging studies in healthy participants have corroborated this general pattern, finding overlapping areas of involvement in each VFT but with contrasts between LFTs and CFTs highlighting greater involvement of frontal regions in the former and of temporal regions in the latter, although the specific regions highlighted are broad ranging and vary considerably across studies.[7,107,108] VFT performance in AD has shown similar dissociations in terms of their associated brain regions. LFTs in these patients have been found to show significant correlations with gray matter density within frontotemporal regions, including the left inferior frontal gyrus, left insula, and bilateral hippocampus and parahippocampal gyrus.[102] As in studies on healthy individuals, areas involved in the LFT were found to overlap with those associated with the semantic task, in this case almost entirely. However, semantic fluency correlated with gray matter in far more extensive areas across the cortex and cerebellum and was uniquely associated with areas of the cingulate gyrus, cerebellum, and fusiform gyrus. As expected, associations between phonemic fluency and gray matter density in this group were limited to areas known to be involved in this function, such as the left inferior frontal gyrus[109] and those reflecting disease severity, such as the hippocampus and parahippocampal gyrus.[110] Semantic fluency however, further highlighted wider areas of structural variation contributing to the semantic component of this task, even when extraneous aspects of VFT performance were controlled for. Similar findings by Vonk et al[25] found a dissociation between areas in which CFT and LFT performances, both at baseline and change over time, correlated with follow-up measures of neurodegeneration in individuals across the AD spectrum. Again, their results found that associations relating to semantic fluency in various AD related areas were unaffected when covarying for baseline LFT performance, whereas the strength of associations between neurodegenerative measures and phonemic fluency were often substantially attenuated when semantic fluency was controlled for. Taken together these findings demonstrate the unique utility of semantic fluency in highlighting areas of cortical degradation in AD populations and suggest that, despite the involvement of executive functions, impairments in LFTs, particularly among those in the early stages of disease, may relate strongly to the word finding dysfunction highlighted best by the CFT.

Discrepancy Scores

Verbal fluency discrepancies have rarely been utilized in imaging studies, perhaps due to the conflicting evidence of their diagnostic efficacy within the literature.[10,70–72,75,78] An early study by Keilp et al,[111] however, did explore the neural correlates of this measure among AD patients and controls using a resting, Xenon-inhalation regional cerebral blood flow measurement (^{133}Xe-rCBF). The advantage of this study was that, unlike previously mentioned research, the verbal fluency discrepancies in this case were calculated using measures of CFT and LFT performance that had been standardized to *z* scores using norms stratified by age, sex (only in LFT) and education levels. As in previous research, this study demonstrated that AD patients performed significantly worse on CFTs relative to LFTs. When correlated with regional cerebral blood flow indices, calculated for parietal and frontal regions, both the standardized CFT and LFT scores were associated with typical AD perfusion deficits within the left parietal cortex. LFT scores were additionally correlated with regional blood flow in frontal regions and CFT scores showing an additional correlation with mean flow,

a measure of diffuse cortical perfusion. Discrepancy scores within the AD group alone were strongly negatively correlated with the frontal index bilaterally, such that the greater the blood flow within these areas, the less discrepancy in the amount of impairment on the two fluency measures. Given the limitations of this study that focused only on indices of the whole frontal and parietal cortices, these findings simply serve to reiterate the notion that more frontally focused cortical damage is likely to be associated with comparable declines in verbal fluency measures[3,4,112–115] and provides little information about the disease mechanisms underlying a decline in discrepancy between verbal fluency measures. More recently, work by our own group looked at the neural correlates of verbal fluency discrepancy scores using whole-brain voxel-based morphometry. Using standardized fluency scores based on normative data from healthy matched controls, discrepancy scores measured in three patient groups, consisting of mildly affected MCI, moderate MCI and AD dementia, were found to correlate with areas of the anterior medial temporal lobes in the mild MCI patients, but correlated more widely with areas of the posterior temporal and occipital cortices in the dementia group. Such findings are in line with a wealth of research suggesting that semantic memory impairments in AD are likely to relate to changes within areas affected by the earliest stages of AD pathology.[25,98,101,116] In later disease stages, however, fluency performance appears to be increasingly influenced by volumetric variance within widespread regions corresponding to the wider semantic network and areas less affected by pathological change, likely reflecting a level of reorganization in the presence of significant medial temporal degradation. As such, it can be concluded that the power of verbal fluency discrepancy scores as a cognitive marker for AD specific cortical degradation, may be greatest in the earliest stages of disease.

Clinical Utility and Differential Diagnosis

Patterns of VFT performance in AD have not only demonstrated value in differentiating AD from healthy aging,[15,17,22–24,69,70] but have further been suggested to have utility in the differential diagnosis of dementia subtypes. Due to the multifactorial nature of fluency performance, the use of CFTs and LFTs in cognitive assessment can reveal specific deficits that may aid in the identification of etiological phenotypes. In frontotemporal lobar degeneration, the behavioral variant of frontotemporal dementia (bvFTD) can be distinguished from the semantic variant using the level of impairment on CFTs. In both cases, individuals show significant impairments on both VFTs when compared with controls.[117] Despite performing similarly on LFTs, however, patients with semantic dementia tend to show significantly greater declines on CFTs when compared with patients with bvFTD, due to significant disruption to functioning of the temporal lobes and degradation of the semantic store.[117,118] These patterns are differentiated from AD by the severity of impairment seen in both tasks.[114] In their early study, Hodges et al[118] outlined patterns of VFT performance in bvFTD, semantic dementia and AD. Crucially, despite performing significantly worse than controls on both VFTs, and worse on the LFT than the AD group, bvFTD patients retained the semantic advantage seen in controls, performing similarly to the AD group on the CFT. In contrast, the semantic dementia group performed significantly worse than AD on both VFTs and were impaired to a greater extent on the CFT than the LFT. Despite the particular utility of CFTs in the diagnosis of AD,[15,17,22–24] use of concurrent LFTs is essential for differential diagnosis. CFT scores alone do very little to explain the underlying mechanisms of their impairment. The differential VFT patterns in FTD groups, however, demonstrate how VFTs can be used in tandem to isolate the underlying pathological processes, namely frontally mediated executive dysfunction among bvFTD, as evidenced by the similar extent of decline on both fluency tasks, vs. temporally mediated impairments in semantic processing in semantic dementia, as evidenced by a significantly greater deficit on CFTs.[117] In AD, disproportionately greater declines in CFTs may, therefore, similarly differentiate patients from those with bvFTD.[114] Conversely, although decline discrepancies between CFT and LFT in AD are similar to those seen in semantic dementia,[119] the severity of impairment in both tasks tends to be significantly greater in semantic dementia patients, therefore allowing for a level of differentiation between the two groups.

Similar findings have supported the use of VFTs in the differentiation of AD from vascular dementia (VaD). Evidence suggests that even in preclinical disease stages, AD and VaD patients may be differentiated by VFT performance due to disproportionately greater declines in LFT in VaD patients contrasting with the semantic impairment characteristic of AD.[120–122] In contrast, the use of VFTs in the differentiation of AD and dementia with Lewy bodies (DLB) is less well

defined. Despite some evidence to suggest that DLB patients may show marginally greater declines on LFTs than AD patients,[123–125] other studies have failed to corroborate such findings.[126] More research is needed in prodromal and preclinical populations to identify whether subtle cognitive deficits due to AD or DLB may manifest in either quantitative or qualitative VFT differences that may aid in differential diagnosis.

Conclusions

In the years since their initial development, VFTs have grown to represent a core component of neuropsychological evaluation, owing to their accessibility and versatility in measuring a range of cognitive functions. In AD, VFTs have proven to be particularly valuable in the identification of semantic processing impairments known to reflect the earliest manifestation of cognitive dysfunction in these individuals.[51] Novel analytical measures that go beyond the standard VFT scoring systems, such as item-based characteristics and VFT performance discrepancies, have been found to perform particularly well in identifying subtle semantic memory decline relating to underlying neuronal changes within rhinal cortices,[64,77] supporting their potential as a cognitive marker for very early neurodegenerative change. Given their already well-established presence in clinical practice, supported by their efficiency and cost effectiveness, the potential value of VFTs in both preclinical screening protocols and as sensitive outcome measures for clinical trials must not be underestimated. Further research aiming to optimize innovative operationalization of VFT scores to utilize best the wealth of potential data available in such simple tasks will, therefore, be instrumental in the move towards earlier diagnosis.

Acknowledgement

MDM is supported by an Alzheimer's Association Research Grant (23AARG-1030190). AV is supported by funding obtained under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.3 - Call for tender No. 341 of 15/03/2022 of Italian Ministry of University and Research funded by the European Union – NextGenerationEU, Project code PE0000006, Concession Decree No. 1553 of 11/10/2022 adopted by the Italian Ministry of University and Research, CUP D93C22000930002, "A multiscale integrated approach to the study of the nervous system in health and disease" (MNESYS).disease" (MNESYS).

Disclosure

The authors report no conflicts of interest.

References

1. Benton AL. Differential behavioral effects in frontal lobe disease. *Neuropsychologia*. 1968;6(1):53–60. doi:10.1016/0028-3932(68)90038-9
2. Newcombe F. Missile wounds of the brain: a study of psychological deficits Oxford University Press ; 1969.
3. Henry JD, Crawford JR. A meta-analytic review of verbal fluency performance in patients with traumatic brain injury. *Neuropsychology*. 2004;18(4):621. doi:10.1037/0894-4105.18.4.621
4. Henry JD, Crawford JR. A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology*. 2004;18 (2):284. doi:10.1037/0894-4105.18.2.284
5. Patterson K, Nestor PJ, Rogers TT. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci*. 2007;8(12):976–987. doi:10.1038/nrn2277
6. Binder JR, Desai RH. The neurobiology of semantic memory. *Trends Cogn Sci*. 2011;15(11):527–536. doi:10.1016/j.tics.2011.10.001
7. Vonk JM, Rizvi B, Lao PJ, et al. Letter and category fluency performance correlates with distinct patterns of cortical thickness in older adults. *Cerebral Cortex*. 2019;29(6):2694–2700. doi:10.1093/cercor/bhy138
8. Reverberi C, Cherubini P, Baldinelli S, Luzzi S. Semantic fluency: cognitive basis and diagnostic performance in focal dementias and Alzheimer's disease. *Cortex*. 2014;54:150–164. doi:10.1016/j.cortex.2014.02.006
9. Mueller KD, Koscik RL, LaRue A, et al. Verbal fluency and early memory decline: results from the Wisconsin registry for Alzheimer's prevention. *Arch Clin Neuropsychol*. 2015;30(5):448–457. doi:10.1093/arclin/acv030
10. Henry JD, Crawford JR, Phillips LH. Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis. *Neuropsychologia*. 2004;42(9):1212–1222. doi:10.1016/j.neuropsychologia.2004.02.001
11. Monsch AU, Bondi MW, Butters N, Salmon DP, Katzman R, Thal LJ. Comparisons of verbal fluency tasks in the detection of dementia of the Alzheimer type. *Arch Neurol*. 1992;49(12):1253–1258. doi:10.1001/archneur.1992.00530360051017
12. Murphy KJ, Rich JB, Troyer AK. Verbal fluency patterns in amnestic mild cognitive impairment are characteristic of Alzheimer's type dementia. *J Int Neuropsychol Soc*. 2006;12(4):570–574. doi:10.1017/S1355617706060590

13. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):270–279. doi:10.1016/j.jalz.2011.03.008
14. Vaughan RM, Coen RF, Kenny R, Lawlor BA. Preservation of the semantic verbal fluency advantage in a large population-based sample: normative data from the TILDA study. *J Int Neuropsychol Soc*. 2016;22(5):570–576. doi:10.1017/S1355617716000291
15. Chasles M-J, Tremblay A, Escudier F, et al. An examination of semantic impairment in amnestic MCI and AD: what can we learn from verbal fluency? *Arch Clin Neuropsychol*. 2020;35(1):22–30. doi:10.1093/arclin/acz018
16. Liampas I, Folia V, Morfakidou R, et al. Language differences among individuals with normal cognition, amnestic and non-amnestic MCI, and Alzheimer's disease. *Arch Clin Neuropsychol*. 2022. doi:10.1093/arclin/acac080
17. Clark LJ, Gatz M, Zheng L, Chen Y-L, McCleary C, Mack WJ. Longitudinal verbal fluency in normal aging, preclinical, and prevalent Alzheimer's disease. *Am J Alzheimers Dis Other Demen*. 2009;24(6):461–468. doi:10.1177/1533317509345154
18. Hedden T, Gabrieli JD. Insights into the ageing mind: a view from cognitive neuroscience. *Nat Rev Neurosci*. 2004;5(2):87–96. doi:10.1038/nrn1323
19. Jack JCR, Knopman DS, Jagust WJ, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol*. 2010;9(1):119–128. doi:10.1016/S1474-4422(09)70299-6
20. García-Herranz S, Díaz-Mardomingo M, Venero C, Peraita H. Accuracy of verbal fluency tests in the discrimination of mild cognitive impairment and probable Alzheimer's disease in older Spanish monolingual individuals. *Aging Neuropsychol Cogn*. 2020;27(6):826–840. doi:10.1080/13825585.2019.1698710
21. Gallucci M, Di Battista ME, Battistella G, Falcone C, Bisiacchi PS, Di Giorgi E. Neuropsychological tools to predict conversion from amnestic mild cognitive impairment to dementia the TREDEM registry. *Aging Neuropsychol Cogn*. 2018;25(4):550–560. doi:10.1080/13825585.2017.1349869
22. Folia V, Liampas I, Siokas V, et al. Language performance as a prognostic factor for developing Alzheimer's clinical syndrome and mild cognitive impairment: results from the population-based HELIAD cohort. *J Int Neuropsychol Soc*. 2022;1–9. doi:10.1017/S1355617722000376
23. Wong CH, Leung GT, Fung AW, Chan WC, Lam LC. Cognitive predictors for five-year conversion to dementia in community-dwelling Chinese older adults. *Int Psychogeriatr*. 2013;25(7):1125–1134. doi:10.1017/S1041610213000161
24. Gustavson DE, Elman JA, Panizzon MS, et al. Association of baseline semantic fluency and progression to mild cognitive impairment in middle-aged men. *Neurology*. 2020;95(8):e973–e983. doi:10.1212/WNL.0000000000010130
25. Vonk JM, Bouteloup V, Mangin JF, et al. Semantic loss marks early Alzheimer's disease-related neurodegeneration in older adults without dementia. *Alzheimers Dement*. 2020;12(1):e12066.
26. Rosen VM, Sunderland T, Levy J, et al. Apolipoprotein E and category fluency: evidence for reduced semantic access in healthy normal controls at risk for developing Alzheimer's disease. *Neuropsychologia*. 2005;43(4):647–658. doi:10.1016/j.neuropsychologia.2004.06.022
27. Papp KV, Mormino EC, Amariglio RE, et al. Biomarker validation of a decline in semantic processing in preclinical Alzheimer's disease. *Neuropsychology*. 2016;30(5):624. doi:10.1037/neu0000246
28. Papp KV, Rentz DM, Orlovsky I, Sperling RA, Mormino EC. Optimizing the preclinical Alzheimer's cognitive composite with semantic processing: the PACC5. *Alzheimers Dement*. 2017;3(4):668–677.
29. Mirandez RM, Aprahamian I, Talib LL, Forlenza OV, Radanovic M. Multiple category verbal fluency in mild cognitive impairment and correlation with CSF biomarkers for Alzheimer's disease. *Int Psychogeriatr*. 2017;29(6):949–958. doi:10.1017/S1041610217000102
30. Amieva H, Le Goff M, Millet X, et al. Prodromal Alzheimer's disease: successive emergence of the clinical symptoms. *Ann Neurol*. 2008;64(5):492–498. doi:10.1002/ana.21509
31. Kempler D, Teng EL, Dick M, Taussig IM, Davis DS. The effects of age, education, and ethnicity on verbal fluency. *J Int Neuropsychol Soc*. 1998;4(6):531–538. doi:10.1017/S1355617798466013
32. Rosselli M, Ardila A, Salvatierra J, Marquez M, Luis M, Weekes VA. A cross-linguistic comparison of verbal fluency tests. *Int J Neurosci*. 2002;112(6):759–776. doi:10.1080/00207450290025752
33. Pekkala S, Goral M, Hyun J, Obler LK, Erkinjuntti T, Albert ML. Semantic verbal fluency in two contrasting languages. *Clin Linguist Phon*. 2009;23(6):431–445. doi:10.1080/02699200902839800
34. Acevedo A, Loewenstein DA, Barker WW, et al. Category fluency test: normative data for English-and Spanish-speaking elderly. *J Int Neuropsychol Soc*. 2000;6(7):760. doi:10.1017/S1355617700677032
35. Garrard P, Ralph MA, Watson PC, Powis J, Patterson K, Hodges JR. Longitudinal profiles of semantic impairment for living and nonliving concepts in dementia of Alzheimer's type. *J Cogn Neurosci*. 2001;13(7):892–909. doi:10.1162/089892901753165818
36. Zannino GD, Perri R, Carlesimo GA, Pasqualetti P, Caltagirone C. Category-specific impairment in patients with Alzheimer's disease as a function of disease severity: a cross-sectional investigation. *Neuropsychologia*. 2002;40(13):2268–2279. doi:10.1016/S0028-3932(02)00110-0
37. Grossman M, Peelle JE, Smith EE, et al. Category-specific semantic memory: converging evidence from bold fMRI and Alzheimer's disease. *Neuroimage*. 2013;68:263–274. doi:10.1016/j.neuroimage.2012.11.057
38. Krumm S, Berres M, Kivisaari SL, et al. Cats and apples: semantic fluency performance for living things identifies patients with very early Alzheimer's disease. *Arch Clin Neuropsychol*. 2021;36(5):838–843. doi:10.1093/arclin/acaa109
39. Capitani E, Laiacona M, Mahon B, Caramazza A. What are the facts of semantic category-specific deficits? A critical review of the clinical evidence. *Cogn Neuropsychol*. 2003;20(3–6):213–261. doi:10.1080/02643290244000266
40. Kwak S, Kim H, Kim H, Youm Y, Chey J. Distributed functional connectivity predicts neuropsychological test performance among older adults. *Hum Brain Mapp*. 2021;42(10):3305–3325. doi:10.1002/hbm.25436
41. Gruenewald PJ, Lockhead GR. The free recall of category examples. *J Exp Psychol*. 1980;6(3):225.
42. Aita SL, Beach JD, Taylor SE, Borgogna NC, Harrell MN, Hill BD. Executive, language, or both? An examination of the construct validity of verbal fluency measures. *Appl Neuropsychol*. 2019;26(5):441–451. doi:10.1080/23279095.2018.1439830
43. Elgamal SA, Roy EA, Sharratt MT. Age and verbal fluency: the mediating effect of speed of processing. *Can Geriatr J*. 2011;14(3):66. doi:10.5770/cgj.v14i3.17
44. Shao Z, Roelofs A, Meyer AS. Predicting naming latencies for action pictures: Dutch norms. *Behav Res Methods*. 2014;46(1):274–283. doi:10.3758/s13428-013-0358-6

45. Greenberg DL, Keane MM, Ryan L, Verfaellie M. Impaired category fluency in medial temporal lobe amnesia: the role of episodic memory. *J Neurosci*. 2009;29(35):10900–10908. doi:10.1523/JNEUROSCI.1202-09.2009
46. Bizzozero I, Scotti S, Clerici F, Pomati S, Laiacona M, Capitani E. On which abilities are category fluency and letter fluency grounded a confirmatory factor analysis of 53 Alzheimer's dementia patients. *Dement Geriatr Cogn Dis Extra*. 2013;3(1):179–191. doi:10.1159/000351418
47. Whiteside DM, Kealey T, Semla M, et al. Verbal fluency: language or executive function measure? *Appl Neuropsychol*. 2016;23(1):29–34. doi:10.1080/23279095.2015.1004574
48. Greenaway MC, Smith GE, Tangalos EG, Geda YE, Ivnik RJ. Mayo older Americans normative studies: factor analysis of an expanded neuropsychological battery. *Clin Neuropsychol*. 2009;23(1):7–20. doi:10.1080/13854040801891686
49. Scheuringer A, Wittig R, Pletzer B. Sex differences in verbal fluency: the role of strategies and instructions. *Cogn Process*. 2017;18(4):407–417. doi:10.1007/s10339-017-0801-1
50. De Marco M, Venneri A. Serial recall order of category fluency words: exploring its neural underpinnings. *Front Psychol*. 2022;12:777838.
51. Venneri A, Jahn-Carta C, De Marco M, Quaranta D, Marra C. Diagnostic and prognostic role of semantic processing in preclinical Alzheimer's disease. *Biomark Med*. 2018;12(6):637–651. doi:10.2217/bmm-2017-0324
52. Murphy DH, Castel AD. Age-related similarities and differences in the components of semantic fluency: analyzing the originality and organization of retrieval from long-term memory. *Aging Neuropsychol Cogn*. 2021;28(5):748–761. doi:10.1080/13825585.2020.1817844
53. Quaranta D, Caprara A, Piccininni C, Vita MG, Gainotti G, Marra C. Standardization, clinical validation, and typicality norms of a new test assessing semantic verbal fluency. *Arch Clin Neuropsychol*. 2016;31(5):434–445. doi:10.1093/arclin/acw034
54. Brysbaert M, Mandera P, Keuleers E. The word frequency effect in word processing: an updated review. *Curr Dir Psychol Sci*. 2018;27(1):45–50. doi:10.1177/0963721417727521
55. Warriner AB, Kuperman V, Brysbaert M. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav Res Methods*. 2013;45(4):1191–1207. doi:10.3758/s13428-012-0314-x
56. Forbes-McKay KE, Ellis AW, Shanks MF, Venneri A. The age of acquisition of words produced in a semantic fluency task can reliably differentiate normal from pathological age related cognitive decline. *Neuropsychologia*. 2005;43(11):1625–1632. doi:10.1016/j.neuropsychologia.2005.01.008
57. Kuperman V, Stadthagen-Gonzalez H, Brysbaert M. Age-of-acquisition ratings for 30,000 English words. *Behav Res Methods*. 2012;44(4):978–990. doi:10.3758/s13428-012-0210-4
58. Wakefield SJ, Blackburn DJ, Harkness K, Khan A, Reuber M, Venneri A. Distinctive neuropsychological profiles differentiate patients with functional memory disorder from patients with amnestic-mild cognitive impairment. *Acta Neuropsychiatr*. 2018;30(2):90–96. doi:10.1017/neu.2017.21
59. Hurks P, Hendriksen J, Vles J, et al. Verbal fluency over time as a measure of automatic and controlled processing in children with ADHD. *Brain Cogn*. 2004;55(3):535–544. doi:10.1016/j.bandc.2004.03.003
60. Sailor KM, Zimmerman ME, Sanders AE. Differential impacts of age of acquisition on letter and semantic fluency in Alzheimer disease patients and healthy older adults. *Q J Exp Psychol*. 2011;64(12):2383–2391. doi:10.1080/17470218.2011.596660
61. Sailor K, Antoine M, Diaz M, Kuslansky G, Kluger A. The effects of Alzheimer's disease on item output in verbal fluency tasks. *Neuropsychology*. 2004;18(2):306–314. doi:10.1037/0894-4105.18.2.306
62. Vita MG, Marra C, Spinelli P, et al. Typicality of words produced on a semantic fluency task in amnesic mild cognitive impairment: linguistic analysis and risk of conversion to dementia. *J Alzheimer's Dis*. 2014;42(4):1171–1178. doi:10.3233/JAD-140570
63. Vonk JMJ, Flores RJ, Rosado D, et al. Semantic network function captured by word frequency in nondemented APOE epsilon4 carriers. *Neuropsychology*. 2019;33(2):256–262. doi:10.1037/neu0000508
64. Venneri A, McGeown WJ, Hietanen HM, Guerrini C, Ellis AW, Shanks MF. The anatomical bases of semantic retrieval deficits in early Alzheimer's disease. Research Support, Non-U.S. Gov't. *Neuropsychologia*. 2008;46(2):497–510. doi:10.1016/j.neuropsychologia.2007.08.026
65. Quaranta D, Piccininni C, Caprara A, Malandrino A, Gainotti G, Marra C. Semantic relations in a categorical verbal fluency test: an exploratory investigation in mild cognitive impairment. *Front Psychol*. 2019;10:2797. doi:10.3389/fpsyg.2019.02797
66. Crowe SF. Decrease in performance on the verbal fluency test as a function of time: evaluation in a young healthy sample. *J Clin Exp Neuropsychol*. 1998;20(3):391–401. doi:10.1076/jcen.20.3.391.810
67. De Marco M, Blackburn DJ, Venneri A. Serial recall order and semantic features of category fluency words to study semantic memory in normal ageing. *Front Aging Neurosci*. 2021;13:678588. doi:10.3389/fnagi.2021.678588
68. Kwak S, Oh DJ, Jeon YJ, et al. Utility of machine learning approach with neuropsychological tests in predicting functional impairment of Alzheimer's disease. *J Alzheimer's Dis*. 2022;85(3):1357–1372. doi:10.3233/JAD-215244
69. Marra C, Piccininni C, Masone Iacobucci G, et al. Semantic memory as an early cognitive marker of Alzheimer's disease: role of category and phonological verbal fluency tasks. *J Alzheimer's Dis*. 2021;81(2):619–627. doi:10.3233/JAD-201452
70. Vaughan RM, Coen RF, Kenny R, Lawlor BA. Semantic and phonemic verbal fluency discrepancy in mild cognitive impairment: potential predictor of progression to Alzheimer's disease. *J Am Geriatr Soc*. 2018;66(4):755–759. doi:10.1111/jgs.15294
71. Teng E, Leone-Friedman J, Lee GJ, et al. Similar verbal fluency patterns in amnestic mild cognitive impairment and Alzheimer's disease. *Arch Clin Neuropsychol*. 2013;28(5):400–410. doi:10.1093/arclin/act039
72. Cerhan JH, Ivnik RJ, Smith GE, Tangalos EC, Petersen RC, Boeve BF. Diagnostic utility of letter fluency, category fluency, and fluency difference scores in Alzheimer's disease. *Clin Neuropsychol*. 2002;16(1):35–42. doi:10.1076/clin.16.1.35.8326
73. Nutter-Upham KE, Saykin AJ, Rabin LA, et al. Verbal fluency performance in amnestic MCI and older adults with cognitive complaints. *Arch Clin Neuropsychol*. 2008;23(3):229–241. doi:10.1016/j.acn.2008.01.005
74. Rinehardt E, Eichstaedt K, Schinka JA, et al. Verbal fluency patterns in mild cognitive impairment and Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2014;38(1–2):1–9. doi:10.1159/000355558
75. Laws KR, Duncan A, Gale TM. 'Normal' semantic-phonemic fluency discrepancy in Alzheimer's disease? A meta-analytic study. *Cortex*. 2010;46(5):595–601. doi:10.1016/j.cortex.2009.04.009
76. Brandt J, Manning KJ. Patterns of word-list generation in mild cognitive impairment and Alzheimer's disease. *Clin Neuropsychol*. 2009;23(5):870–879. doi:10.1080/13854040802585063

77. Wright LM, De Marco M, Venneri A. Verbal fluency discrepancies as a marker of the prehippocampal stages of Alzheimer's disease. *Neuropsychology*. 2022. doi:10.1037/neu0000836
78. Lonie JA, Herrmann LL, Tierney KM, et al. Lexical and semantic fluency discrepancy scores in aMCI and early Alzheimer's disease. *J Neuropsychol*. 2009;3:79–92. doi:10.1348/174866408x289935
79. Salmon DP, Heindel WC, Lange KL. Differential decline in word generation from phonemic and semantic categories during the course of Alzheimer's disease: implications for the integrity of semantic memory. *J Int Neuropsychol Soc*. 1999;5(7):692–703. doi:10.1017/S1355617799577126
80. Troyer AK, Moscovitch M, Winocur G. Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology*. 1997;11(1):138. doi:10.1037/0894-4105.11.1.138
81. Troyer AK, Moscovitch M, Winocur G, Alexander MP, Stuss D. Clustering and switching on verbal fluency: the effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia*. 1998a;36:449–504. doi:10.1016/S0028-3932(97)00152-8
82. Troyer AK, Moscovitch M, Winocur G, Leach L, Freedman M. Clustering and switching on verbal fluency tests in Alzheimer's and *Parkinson*'s disease. *J Int Neuropsychol Soc*. 1998b;4:137–143. doi:10.1017/S1355617798001374
83. Gomez RG, White DA. Using verbal fluency to detect very mild dementia of the Alzheimer type. Archives of Clinical. *Neuropsychology*. 2006;21(8):771–775.
84. Weakley A, Schmitter-Edgecombe M. Analysis of verbal fluency ability in Alzheimer's disease: the role of clustering, switching and semantic proximities. Archives of Clinical. *Neuropsychology*. 2014;29(3):256–268.
85. Fagundo AB, López S, Romero M, Guarch J, Marcos T, Salamero M. Clustering and switching in semantic fluency: predictors of the development of Alzheimer's disease. *Int J Geriatr Psychiatry*. 2008;23(10):1007–1013. doi:10.1002/gps.2025
86. Raoux N, Amieva H, Le Goff M, et al. Clustering and switching processes in semantic verbal fluency in the course of Alzheimer's disease subjects: results from the PAQUID longitudinal study. *Cortex*. 2008;44(9):1188–1196. doi:10.1016/j.cortex.2007.08.019
87. Pekkala S, Albert ML, Spiro III A, Erkinjuntti TI. Perseveration in Alzheimer's disease. *Dement Geriatr Cogn Disord*. 2008;25(2):109–114. doi:10.1159/000112476
88. Pakhomov SV, Eberly LE, Knopman DS. Recurrent perseverations on semantic verbal fluency tasks as an early marker of cognitive impairment. *J Clin Exp Neuropsychol*. 2018;40(8):832–840. doi:10.1080/13803395.2018.1438372
89. Liampas I, Folia V, Zoupa E, et al. Qualitative verbal fluency components as prognostic factors for developing Alzheimer's dementia and mild cognitive impairment: results from the population-based HELIAD Cohort. *Medicina*. 2022;58(12):1814. doi:10.3390/medicina58121814
90. Giffard B, Laisney M, Mezenge F, de la Sayette V, Eustache F, Desgranges B. The neural substrates of semantic memory deficits in early Alzheimer's disease: clues from semantic priming effects and FDG-PET. *Neuropsychologia*. 2008;46(6):1657–1666. doi:10.1016/j.neuropsychologia.2007.12.031
91. Carter SF, Embleton KV, Anton-Rodriguez JM, Burns A, Ralph MAL, Herholz K. Regional neuronal network failure and cognition in late-onset sporadic Alzheimer disease. *Am J Neuroradiol*. 2014;35(6):S18–S30. doi:10.3174/ajnr.A3895
92. Apostolova LG, Lu P, Rogers S, et al. 3D mapping of language networks in clinical and pre-clinical Alzheimer's disease. Research Support, N.I. H., Extramural Research Support, Non-U.S. Gov't. *Brain Lang*. 2008;104(1):33–41. doi:10.1016/j.bandl.2007.03.008
93. Dos Santos V, Thomann PA, Wustenberg T, Seidl U, Essig M, Schroder J. Morphological cerebral correlates of CERAD test performance in mild cognitive impairment and Alzheimer's disease. *J Alzheimer's Dis*. 2011;23(3):411–420. doi:10.3233/JAD-2010-100156
94. Eastman JA, Hwang KS, Lazaris A, et al. Cortical thickness and semantic fluency in Alzheimer's disease and mild cognitive impairment. *Am J Alzheimers Dis*. 2013;1(2):81–92.
95. McDonald CR, Gharapetian L, McEvoy LK, et al. Relationship between regional atrophy rates and cognitive decline in mild cognitive impairment. *Neurobiol Aging*. 2012;33(2):242–253. doi:10.1016/j.neurobiolaging.2010.03.015
96. Peter J, Kaiser J, Landerer V, et al. Category and design fluency in mild cognitive impairment: performance, strategy use, and neural correlates. *Neuropsychologia*. 2016;93(Pt A):21–29.
97. Loewenstein DA, Curiel RE, Wright C, et al. Recovery from proactive semantic interference in mild cognitive impairment and normal aging: relationship to atrophy in brain regions vulnerable to Alzheimer's disease. Research Support, N.I.H., Extramural. *J Alzheimer's Dis*. 2017;56(3):1119–1126. doi:10.3233/JAD-160881
98. Venneri A, Mitolo M, Beltrachini L, et al. Beyond episodic memory: semantic processing as independent predictor of hippocampal/perirhinal volume in aging and mild cognitive impairment due to Alzheimer's disease. *Neuropsychology*. 2019;33(4):523–533. doi:10.1037/neu0000534
99. Meyer P, Feldkamp H, Hoppstädter M, et al. Using voxel-based morphometry to examine the relationship between regional brain volumes and memory performance in amnestic mild cognitive impairment. *Front Behav Neurosci*. 2013;7:89. doi:10.3389/fnbeh.2013.00089
100. Hirni DI, Kivisaari SL, Krumm S, et al. Neuropsychological markers of medial perirhinal and entorhinal cortex functioning are impaired twelve years preceding diagnosis of Alzheimer's dementia. Research Support, Non-U.S. Gov't. *J Alzheimer's Dis*. 2016;52(2):573–580. doi:10.3233/JAD-150158
101. Hirni DI, Kivisaari SL, Monsch AU, Taylor KI. Distinct neuroanatomical bases of episodic and semantic memory performance in Alzheimer's disease. Research Support, Non-U.S. Gov't. *Neuropsychologia*. 2013;51(5):930–937. doi:10.1016/j.neuropsychologia.2013.01.013
102. Rodríguez-Aranda C, Waterloo K, Johnsen SH, et al. Neuroanatomical correlates of verbal fluency in early Alzheimer's disease and normal aging. *Brain Lang*. 2016;155:24–35. doi:10.1016/j.bandl.2016.03.001
103. Friederici AD, Gierhan SM. The language network. *Curr Opin Neurobiol*. 2013;23(2):250–254. doi:10.1016/j.conb.2012.10.002
104. Hoffman P, Morcom AM. Age-related changes in the neural networks supporting semantic cognition: a meta-analysis of 47 functional neuroimaging studies. *Neurosci Biobehav Rev*. 2018;84:134–150. doi:10.1016/j.neubiorev.2017.11.010
105. Schmidt CSM, Nitschke K, Bormann T, et al. Dissociating frontal and temporal correlates of phonological and semantic fluency in a large sample of left hemisphere stroke patients. *Neuroimage Clin*. 2019;23:101840. doi:10.1016/j.nicl.2019.101840
106. Biesbroek JM, Lim JS, Weaver NA, et al. Anatomy of phonemic and semantic fluency: a lesion and disconnectome study in 1231 stroke patients. *Cortex*. 2021;143:148–163. doi:10.1016/j.cortex.2021.06.019
107. Grogan A, Green DW, Ali N, Crinion JT, Price CJ. Structural correlates of semantic and phonemic fluency ability in first and second languages. *Cerebral Cortex*. 2009;19(11):2690–2698. doi:10.1093/cercor/bhp023

108. Birn RM, Kenworthy L, Case L, et al. Neural systems supporting lexical search guided by letter and semantic category cues: a self-paced overt response fMRI study of verbal fluency. *Neuroimage*. 2010;49(1):1099–1107. doi:10.1016/j.neuroimage.2009.07.036
109. Costafreda SG, Fu CH, Lee L, Everitt B, Brammer MJ, David AS. A systematic review and quantitative appraisal of fMRI studies of verbal fluency: role of the left inferior frontal gyrus. *Hum Brain Mapp*. 2006;27(10):799–810. doi:10.1002/hbm.20221
110. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol*. 1991;82(4):239–259. doi:10.1007/BF00308809
111. Keilp J, Gorlyn M, Alexander G, Stern Y, Prohovnik I. Cerebral blood flow patterns underlying the differential impairment in category vs letter fluency in Alzheimer's disease. *Neuropsychologia*. 1999;37(11):1251–1261. doi:10.1016/S0028-3932(99)00032-9
112. Kitabayashi Y, Ueda H, Tsuchida H, et al. Relationship between regional cerebral blood flow and verbal fluency in Alzheimer's disease. *Psychiatry Clin Neurosci*. 2001;55(5):459–463. doi:10.1046/j.1440-1819.2001.00890.x
113. Mummery CJ, Patterson K, Hodges JR, Wise RJ. Generating 'tiger'as an animal name or a word beginning with T: differences in brain activation. *Proc Royal Soc B*. 1996;263(1373):989–995.
114. Rascovsky K, Salmon DP, Hansen LA, Thal LJ, Galasko D. Disparate letter and semantic category fluency deficits in autopsy-confirmed frontotemporal dementia and Alzheimer's disease. *Neuropsychology*. 2007;21(1):20. doi:10.1037/0894-4105.21.1.20
115. Capitani E, Rosci C, Saetti MC, Laiacona M. Mirror asymmetry of category and letter fluency in traumatic brain injury and Alzheimer's patients. *Neuropsychologia*. 2009;47(2):423–429. doi:10.1016/j.neuropsychologia.2008.09.016
116. Barbeau EJ, Didic M, Joubert S, et al. Extent and neural basis of semantic memory impairment in mild cognitive impairment. *J Alzheimer's Dis*. 2012;28(4):823–837. doi:10.3233/JAD-2011-110989
117. Laisney M, Matuszewski V, Mézenge F, et al. The underlying mechanisms of verbal fluency deficit in frontotemporal dementia and semantic dementia. *J Neurol*. 2009;256:1083–1094. doi:10.1007/s00415-009-5073-y
118. Hodges JR, Patterson K, Ward R, et al. The differentiation of semantic dementia and frontal lobe dementia (temporal and frontal variants of frontotemporal dementia) from early Alzheimer's disease: a comparative neuropsychological study. *Neuropsychology*. 1999;13(1):31.l. doi:10.1037/0894-4105.13.1.31
119. Rogers TT, Ivanoiu A, Patterson K, Hodges JR. Semantic memory in Alzheimer's disease and the frontotemporal dementias: a longitudinal study of 236 patients. *Neuropsychology*. 2006;20(3):319. doi:10.1037/0894-4105.20.3.319
120. Duff-Canning S, Leach L, Stuss D, Ngo L, Black S. Diagnostic utility of abbreviated fluency measures in Alzheimer disease and vascular dementia. *Neurology*. 2004;62(4):556–562. doi:10.1212/WNL.62.4.556
121. Jones S, Laukka EJ, Bäckman L. Differential verbal fluency deficits in the preclinical stages of Alzheimer's disease and vascular dementia. *Cortex*. 2006;42(3):347–355. doi:10.1016/S0010-9452(08)70361-7
122. Ramirez-Gomez L, Zheng L, Reed B, et al. Neuropsychological profiles differentiate Alzheimer disease from subcortical ischemic vascular dementia in an autopsy-defined cohort. *Dement Geriatr Cogn Disord*. 2017;44(1–2):1. doi:10.1159/000477344
123. Hansen L, Salmon D, Galasko D, et al. The Lewy body variant of Alzheimer's disease: a clinical and pathological entity. *Neurology*. 1990;40:1–8. doi:10.1212/WNL.40.1.1
124. Calderon J, Perry RJ, Erzinclioglu SW, Berrios GE, Dening T, Hodges JR. Perception, attention, and working memory are disproportionately impaired in dementia with Lewy bodies compared with Alzheimer's disease. *J Neurol Neurosurg Psychiatry*. 2001;70(2):157–164. doi:10.1136/jnnp.70.2.157
125. Lambon-Ralph M, Powell J, Howard D, Whitworth AB, Garrard P, Hodges JR. Semantic memory is impaired in both dementia with Lewy bodies and dementia of Alzheimer's type: a comparative neuropsychological study and literature review. *J Neurol Neurosurg Psychiatry*. 2020;70(2):149–156. doi:10.1136/jnnp.70.2.149
126. Noe E, Marder K, Bell KL, Jacobs DM, Manly JJ, Stern Y. Comparison of dementia with Lewy bodies to Alzheimer's disease and Parkinson's disease with dementia. *Mov Disorders*. 2004;19(1):60–67. doi:10.1002/mds.10633