Clustering and switching during a semantic verbal fluency test contribute to differential diagnosis of cognitive impairment

Qianhua Zhao, Qihao Guo, Zhen Hong

Department of Neurology, Huashan Hospital, Fudan University, Shanghai 200040, China

Corresponding author: Qihao Guo. E-mail: dr.guqihao@126.com

© Shanghai Institutes for Biological Sciences, CAS and Springer-Verlag Berlin Heidelberg 2013

Abstract

The verbal fluency test (VFT) can be dissociated into “clustering” (generating words within subcategories) and “switching” (shifting between clusters), which may be valuable in differential diagnosis. In the current study, we investigated the validity of VFT in the differential diagnosis of Alzheimer’s disease (AD, n = 65), vascular dementia (VaD, n = 65), mild cognitive impairment (MCI, n = 92), and vascular cognitive impairment without dementia (VCIND, n = 76) relative to cognitively normal senior controls (NC, n = 374). We found that in the NC group, the total correct score was significantly correlated with age and education; males generated more subcategories; cluster size increased with education, and subcategory and switching decreased with age. A significantly progressive advantage was observed in VFT scores in the sequence NC > MCI/VCIND > AD/VaD, and this significantly discriminated dementia patients from the other groups. AD patients performed better in all four VFT scores than VaD patients. Subcategory and switching scores significantly distinguished AD from VaD patients (AD > VaD; mean difference, 0.50 for subcategory, \( P < 0.05 \); 0.71 for switching, \( P < 0.05 \)). MCI patients scored higher than VCIND patients, but the difference did not reach statistical significance. These results suggest that semantic VFT is useful for the detection of MCI and VCIND, and in the differential diagnosis of cognitive impairment.

Keywords: Alzheimer’s disease; dementia; verbal fluency; mild cognitive impairment

INTRODUCTION

The verbal fluency test (VFT)\(^1\) is a widely-used neuropsychological test of language production, requiring participants to generate words in response to a cue. This requires the retrieval of lexically-associated words from long-term memory and involves high demands on frontally-mediated strategic processes. The purpose of the VFT is to evaluate the spontaneous production of words beginning with a given letter or belonging to a given class in a limited amount of time. For category (semantic association) fluency, the participant is asked to produce as many items (e.g., animal, supermarket item, fruit, vegetable) as quickly as possible (in 60 or 30 s)\(^1\). The VFT is widely used in the diagnosis, efficacy assessment, and prognosis of many neurological disorders such as dementia, epilepsy, head trauma, central nervous system infection, and mental retardation.

The VFT is time-saving (takes only a few minutes) and provides valuable information about semantic memory, executive function, naming, attention-shifting and sequencing. These functions involve cortical and subcortical structures, such as the frontal lobe, temporal lobe, and striatum\(^2\). The total number of correct words generated is most commonly used as the index of task performance. However, this index misses much of the information, and does not perform well in discriminating between dementia and cognitive impairment\(^3\). In addition to the total number of correct words generated, a qualitative analysis of the word sequence provides valuable information on the patient’s impaired cognitive components. Troyer, Moscovich and Winocur isolated two main components in VFT tasks: “clustering”, the ability...
to produce words within phonetic and semantic subcategories; and “switching”, the ability to shift between clusters\[3\]. Clustering relies more on temporal lobe functions such as word storage and semantic memory, while switching relies more on frontal lobe functions such as sorting and attention-shifting\[3,4\].

Previous studies reported that the neuropsychological deficits of Alzheimer’s disease (AD) are characterized by a breakdown in the organization and structure of semantic knowledge and impairments in executive function\[5\]. Vascular dementia (VaD), the second most common cause of dementia, is characterized by subcortical dysfunction\[6\]. Whether the VFT can differentiate between subtypes of dementia has been widely studied\[5-8\]. Canning reported that the total correct number of responses on the letter fluency test was helpful in differentiating VaD from AD in a relatively small sample\[9\], but this test cannot be directly used in a Mandarin-speaking Chinese population. Besides, few studies have investigated whether clustering and switching can discriminate between AD and VaD.

Mild cognitive impairment (MCI) and vascular cognitive impairment without dementia (VCIND) have been reported to represent preclinical stages and are potential treatment targets\[10,11\]. Murphy et al. reported that patterns of verbal fluency performance in amnestic MCI patients are characteristic of AD\[12\]. Thus, it remains unclear whether the VFT and its clustering and switching components are useful in distinguishing between AD and VaD, and between MCI and VCIND.

In the current study, we used a semantic VFT rather than phonemic fluency, because it is easier for participants to generate clusters. As such, the semantic VFT is suitable for analyzing clustering and switching. “Supermarket” was chosen as the defined category, since senior individuals in China typically have experience with shopping in supermarkets, regardless of gender, educational level or other aspects of their background.

We administered the semantic VFT using the “supermarket” category (items in a supermarket) to Chinese participants with normal cognition and patients with AD, VaD, MCI, or VCIND. Next, the total correct score and subscores were analyzed to investigate their correlation with demographic variables, and to evaluate their validity in the differential diagnosis of cognitive impairment.

METHODS

Participants

Five groups of participants were recruited: patients with AD, VaD, MCI, VCIND, and cognitively normal controls (NC). Participants were consecutively selected from the Memory Clinic of Huashan Hospital from 2008 to 2010. Normal controls were obtained by random group sampling from the community. The inclusion criteria were: (1) fluency in Mandarin, (2) aged between 55 and 85 years, (3) underwent physical, neurological, and neuropsychological evaluations, (4) underwent neuroimaging (magnetic resonance imaging or computed tomography) when indicated; and (5) good compliance and written informed consent. The exclusion criteria were: (1) diagnosis of depression or anxiety in the past month, and (2) having visual or hearing problems. All participants exhibited good compliance and cooperated well during the cognitive assessment. The study was approved by the Independent Review Board of Huashan Hospital.

The AD group consisted of patients with a diagnosis of probable AD based on the National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer’s Disease and Related Disorders Association criteria\[13\]. Patients in the VaD group were diagnosed according to the guidelines of the National Institute of Neurological Disorders and Stroke and the Association Internationale pour la Recherche et l’Enseignement en Neurosciences\[14\]. The MCI group consisted of both amnestic and non-amnestic MCI patients using the following Petersen criteria\[15,16\]: self- and/or informant-reported cognitive decline at least 3 months in duration; objective evidence of cognitive decline (scored below −1.5 SD of the mean in neuropsychological tests); preserved basic activities of daily living; and no dementia. Diagnosis was based on consensus meetings involving neurologists, neuropsychologists and neuroradiologists. Patients in the VCIND group were diagnosed based on the guidelines of the National Institute of Neurological Disorders and Stroke-Canadian Stroke Network Vascular Cognitive Impairment Harmonization Standards\[17,18\].

Category VFT

All participants took the semantic VFT using “supermarket” as the defined category\[11\]. Participants were asked to pro-
duce as many words as possible after receiving the following instructions: “I am going to tell you the names of items found in a supermarket: rice, clothes, and books. Can you think of any other items?” The participant then named other items and the examiner corrected them if they produced incorrect responses, then repeated the instructions. Participants were then instructed: “Now, name as many items that are found in a supermarket as possible, as quickly as possible.” Participants were allowed 1 minute (by stopwatch) for the test. If participants stopped before the allotted time, they were encouraged to name more items. The instructions were repeated and hints were given if there was a pause of 15 seconds or more. Timing started immediately after the instructions were given, allowing for extra time when instructions were repeated. The actual words were written down in the order of their production.

This study adopted Troyer’s scoring method\(^3\), resulting in four scores: (1) the sum of all admissible words. Slang terms and foreign words were accepted, but inadmissible words were excluded; (2) clustering: the ability to produce words within semantic subcategories. Clustering refers to the average number of words produced within each cluster or subcategory; (3) subcategory: the number of different types of clusters generated; and (4) switching: the ability to shift between clusters (i.e., the number of shifts). Errors and repetitions were included in the calculation of cluster size and switching, because any word produced provided information about the underlying cognitive process, regardless of whether or not it contributed to the total correct number of words generated\(^3\). If the participant went back to a previous subcategory, the subsequent items generated were recalculated as a new cluster and were not included in the previous subcategory.

In a pilot study, we used a non-arbitrary classification of supermarket items using discrete subcategories. In that study, 100 volunteers were asked to generate supermarket items within 60 s. The words were then grouped according to semantic proximity. Using these data, the supermarket clusters were defined as follows: food, beverages, fruit, vegetables, household items, books and office supplies, clothes and accessories. Items that did not fit the predefined subcategories were classified in the "other" subcategory. Then the analysis was based on these eight subcategories.

### Other Neuropsychological Tests

A battery of tests was administered to all participants, including the cognitively normal controls, to assist diagnosis. The tests were: the Mini-Mental Status Examination (MMSE)\(^19\); the Mattis Dementia Rating Scale\(^20,21\); the Rey Auditory Verbal Learning Test\(^22\); the WMS Logic Memory Test\(^23\); the Rey-Osterrieth Complex Figure Test\(^24\); the Stroop Color and Word Test\(^25\); the Trail Making Test\(^26\); the Similarity Test\(^1\); the Boston Naming Test\(^27\); the Center for Epidemiological Studies-Depression Scale\(^28\); the Activities of Daily Life Scale\(^29\); and the Clinical Dementia Rating Scale\(^30\). The neurologists were blinded to the VFT scores when making a diagnosis of AD, VaD, MCI or VCIND.

### Statistical Analysis

SPSS for Windows 13.0 (SPSS Inc., Chicago, IL) was used for data analysis. Gender, age, education, VFT, and MMSE scores were compared using the \(\chi^2\) test or ANOVA as appropriate. A linear partial correlation coefficient was used to determine associations between VFT scores and demographic indicators. ANCOVA was used to compare variables among five groups and between each pair of groups with age, gender and education as covariates. In comparisons between pairs of groups (AD versus VaD, MCI versus VCIND) the MMSE scores were also adjusted besides the demographic variables. This was an exploratory study that warranted maintaining a relatively liberal \(P\) value, so significance was set at \(P < 0.05\) throughout the analysis.

### RESULTS

#### Demographic Information

The demographic information of the participants is presented in Table 1. The MMSE scores in the AD and VaD groups had a relatively larger range. However, there was no significant difference between AD and VaD patients, or between MCI and VCIND patients in MMSE scores, or any other demographic variables.

#### Correlation of Demographic Variables with VFT Subscores in Cognitively Normal Subjects

Since age, gender and education were correlated in this Chinese Mandarin-speaking senior population, a linear
regression was performed to adjust for the interaction between each demographic variable and to analyze the correlation of each VFT sub-score with the demographic variables in the NC group. The total correct score was found to correlate closely with age and education. Males generated more subcategories than females. Cluster size increased with level of education. Subcategory and switching decreased with age. The results are shown in Table 2.

### VFT in Diagnostic Groups

There was a significant progressive advantage in supermarket-related word fluency in the sequence NC > MCI/VCIND > AD/VaD in the total score and all sub-scores. The ANCOVA for each pair of groups revealed that all four scores (total correctness, subcategory, cluster size, and switching) significantly discriminated dementia from NC ($P < 0.01$), and differentiated MCI/VCIND from dementia ($P < 0.01$). All VFT scores except cluster size distinguished MCI/VCIND from NC ($P < 0.01$, Table 3).

As shown in Table 1, the MMSE score did not differ between AD and VaD, or between MCI and VCIND, indicating that these groups had the same severity of cognitive impairment. But AD patients performed better than VaD pa-

---

### Table 1. Demographic variables in each group (mean ± SD)

<table>
<thead>
<tr>
<th></th>
<th>AD ($n = 65$)</th>
<th>VaD ($n = 65$)</th>
<th>MCI ($n = 92$)</th>
<th>VCIND ($n = 76$)</th>
<th>NC ($n = 374$)</th>
<th>MCI vs VCIND</th>
<th>AD vs VaD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (M/F)</td>
<td>42/23</td>
<td>41/24</td>
<td>55/37</td>
<td>47/29</td>
<td>156/218</td>
<td>0.874</td>
<td>0.742</td>
</tr>
<tr>
<td>Age (years)</td>
<td>72.29 ± 9.54</td>
<td>71.66 ± 9.54</td>
<td>67.55 ± 6.26</td>
<td>64.74 ± 7.56</td>
<td>66.26 ± 7.16</td>
<td>0.099</td>
<td>0.650</td>
</tr>
<tr>
<td>Education</td>
<td>12.66 ± 2.43</td>
<td>12.75 ± 2.46</td>
<td>12.45 ± 2.96</td>
<td>12.51 ± 2.74</td>
<td>12.72 ± 2.79</td>
<td>0.533</td>
<td>0.259</td>
</tr>
<tr>
<td>MMSE</td>
<td>18.23 ± 3.91</td>
<td>19.00 ± 5.21</td>
<td>27.12 ± 1.43</td>
<td>27.14 ± 1.98</td>
<td>27.75 ± 2.00</td>
<td>0.820</td>
<td>0.101</td>
</tr>
</tbody>
</table>

AD, Alzheimer’s disease; F, female; M, male; MCI, mild cognitive impairment; NC, cognitively normal controls; VaD, vascular dementia; VCIND, vascular cognitive impairment non-dementia.

### Table 2. Correlations between demographic variables and verbal fluency test scores in cognitively normal subjects

<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Gender</th>
<th>Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total correct</td>
<td>−0.179**</td>
<td>−0.339</td>
<td>0.297**</td>
</tr>
<tr>
<td>Subcategory</td>
<td>−0.032**</td>
<td>−0.300*</td>
<td>0.000</td>
</tr>
<tr>
<td>Cluster size</td>
<td>−0.027</td>
<td>−0.316</td>
<td>0.129**</td>
</tr>
<tr>
<td>Switching</td>
<td>−0.084**</td>
<td>−0.581</td>
<td>−0.012</td>
</tr>
</tbody>
</table>

Correlation coefficient B values are presented. *$P < 0.05$, **$P < 0.01$.

*The subcategory score was correlated with gender in cognitively normal subjects. Gender-specific subcategory score: men > women (4.13 ± 1.30 vs 3.97 ± 1.25).

### Table 3. Verbal fluency test scores in each group

<table>
<thead>
<tr>
<th>Index</th>
<th>AD ($n = 65$)</th>
<th>VaD ($n = 65$)</th>
<th>MCI ($n = 92$)</th>
<th>VCIND ($n = 76$)</th>
<th>NC ($n = 374$)</th>
<th>Mean difference</th>
<th>P</th>
<th>Mean difference</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AD - VaD</td>
<td></td>
<td>MCI - VCIND</td>
<td></td>
</tr>
<tr>
<td>Total correct</td>
<td>6.68 ± 5.46</td>
<td>6.34 ± 5.23</td>
<td>15.86 ± 6.29</td>
<td>15.71 ± 5.64</td>
<td>20.19 ± 5.83</td>
<td>0.34</td>
<td>0.089</td>
<td>0.15</td>
<td>0.230</td>
</tr>
<tr>
<td>Subcategory</td>
<td>2.72 ± 1.52</td>
<td>2.22 ± 1.50</td>
<td>3.74 ± 1.28</td>
<td>3.64 ± 1.38</td>
<td>4.04 ± 1.27</td>
<td>0.50</td>
<td>0.035</td>
<td>0.10</td>
<td>0.665</td>
</tr>
<tr>
<td>Cluster size</td>
<td>3.30 ± 2.45</td>
<td>2.98 ± 2.54</td>
<td>4.69 ± 1.32</td>
<td>4.64 ± 2.61</td>
<td>4.80 ± 2.48</td>
<td>0.32</td>
<td>0.501</td>
<td>0.05</td>
<td>0.881</td>
</tr>
<tr>
<td>Switching</td>
<td>2.69 ± 2.13</td>
<td>1.98 ± 2.00</td>
<td>4.74 ± 2.78</td>
<td>4.06 ± 2.19</td>
<td>5.94 ± 2.91</td>
<td>0.71</td>
<td>0.036</td>
<td>0.68</td>
<td>0.094</td>
</tr>
</tbody>
</table>

AD, Alzheimer’s disease; MCI, mild cognitive impairment; NC, cognitively normal controls; VaD, vascular dementia; VCIND, vascular cognitive impairment non-dementia. The total correct index was adjusted by covariance analysis for comparison in subcategory, cluster size and switching.
patients in all four VFT scores (Table 3). The subcategory and switching scores significantly distinguished AD from VaD (AD > VaD, mean difference: 0.50 for subcategory, P <0.05; 0.71 for switching, P <0.05). MCI patients performed better than VCIND patients on all four scores, but the difference did not reach significance.

**DISCUSSION**

VFT is widely used in clinical and experimental settings in patients with cognitive impairment[6]. Previous studies have suggested that the category fluency task, rather than phonemic fluency, relies more on semantic associations and is therefore better for distinguishing between AD and VaD[3,4]. Because of language differences, we used the semantic fluency test in the current study.

It has been reported that demographic variables affect verbal fluency[1,31], but this has not been studied in a Mandarin-speaking population. Here, we demonstrated that participants’ performance on the task (total number of correct responses) correlated with both age and education in cognitively normal subjects. Analysis of sub-scores revealed that cluster size was positively correlated with educational level, while switching score correlated negatively with age, consistent with the findings of Lanting et al.[32]. Troyer et al. reported that cluster size reflects semantic knowledge storage, which involves the temporal lobe[3]. It is possible that education helps to enrich an individual’s semantic knowledge, therefore increasing the cluster size. However, switching relies more on the frontal lobe. Since executive function deteriorates over time, it is not surprising that the subcategory and switching scores decreased with age.

The total number of correct words generated is the most commonly used index of performance in fluency tasks. However, this index neither provides information about the cognitive components underlying fluency performance nor necessarily discriminates between different cognitively-impaired populations[6]. Troyer et al. proposed that clustering and switching are the most important aspects of fluency performance[3], and can be used to discriminate Parkinson’s disease dementia and Huntington’s disease from AD[4].

MCI, especially in amnestic MCI patients, is a high-risk factor for AD and is characterized by a selective decline in episodic memory[19]. In addition, VCIND is thought to represent the preclinical state of VaD[17]. Accurately differentiating these conditions is important because different pharmaceutical strategies can modify the course of each disease. Many studies have concerned the differential validity of various neuropsychological tests. Oguro et al. reported that mental flexibility (measured by phonological verbal fluency) can significantly discriminate individuals with VaD from AD patients and normal controls[6]. Jones et al. found that patients with preclinical AD and VaD are similarly impaired in letter fluency, but significantly differed in category fluency[33]. Nutter-Upham et al. found that both phonemic and semantic fluency performance is significantly decreased in amnestic MCI patients compared to cognitively intact older adults, indicating subtle changes in the quality of semantic storage and slowing of retrieval[30]. Murphy et al. identified a progressive advantage in semantic fluency in the sequence, controls > aMCI > AD[12].

It is widely accepted that executive function in patients with VaD tends to be disproportionately impaired, including planning and sequencing, speed of mental processing, performance on unstructured tasks, and attention[7,11,33-35]. Language production may be impaired in patients with VaD, but primary language functions tend to be preserved. Compared to individuals with VaD, patients with AD may exhibit greater deficits in functions (including memory) mediated by posterior cortical structures such as the temporal and parietal lobes. AD patients exhibit a faster rate of information decay, a reduced ability to benefit from cues to facilitate retrieval, and a higher frequency of intrusion errors. In addition, impairment in naming function may exacerbate the deficits in verbal memory tasks.

Many studies have investigated the neuropsychological profiles of AD and VaD. Few, however, have investigated the verbal fluency sub-scores in different cognitively-impaired groups, such as MCI, VCIND, AD, and VaD[11,33]. Since the VFT sub-scores for clustering and switching have been reported to correlate with the temporal and frontal lobes[3], respectively, it is necessary to test the differential validity of those sub-scores between AD and VaD. It would be useful for future studies to determine whether this difference exists in the prodromal state in MCI and VCIND patients.
The current results are in accord with previous results showing that different scores distinguish cognitively-impaired groups in a stepwise manner (NC > MCI > AD; NC > VCIND > VaD)\(^{[12]}\). MCI patients exhibit similar but less severe impairments than AD patients, as do VCIND patients compared to VaD patients. The current findings also indicated that semantic memory deficits, which are a characteristic change in AD, also existed in MCI patients. In contrast, executive dysfunction is a common impairment among both VCIND and VaD patients, and is thought to be attributable to subcortical deficits.

In this study, the total correct number of responses in the supermarket-related word fluency test failed to discriminate AD from VaD patients and MCI from VCIND patients. However, the subcategory and switching sub-scores successfully distinguished AD from VaD patients, and switching sub-score showed a similar trend between MCI and VCIND patients. Patients with vascular disorders (VCIND and VaD) were found to switch less than patients with MCI and AD. It is known that the cognitive deficits in VaD and VCIND patients are characterized by a disproportionate disturbance in frontal executive function, while switching relies mostly on the frontal lobe. This suggests that subcategory and switching may provide a suitable candidate index for differentiating AD and VaD.

We did not include mixed dementia in the current study because the condition is controversial, and there is no consensus on the diagnostic criteria. In addition, mixed dementia patients exhibit both degenerative and vascular pathology, which could confuse the differential validity of the VFT. Moreover, we failed to find a significant difference between MCI and VCIND in all four verbal fluency indices. This result may have arisen because we included both amnestic and non-amnestic MCI patients. Although amnestic MCI is regarded as the prodromal state of AD and involves memory loss as a characteristic impairment, executive function may also be affected in amnestic MCI patients. Non-amnestic MCI involves non-memory cognitive impairments, particularly in executive function. Since a considerable proportion of MCI patients also exhibit executive function impairments, sub-scores reflecting frontal lobe function, such as switching and subcategory in verbal fluency, may lose validity in differentiating between MCI and VCIND.

Our study has several advantages. Clustering and switching have not been previously described in a supermarket-related word fluency task. We chose “supermarket” as the category because it is familiar to many people, regardless of the demographic characteristics and educational background. In addition, this was the first reported comparison in a Mandarin-speaking population. Thus, the results indicate that clustering and switching in the VFT are not limited to English-speaking individuals, but generalize to speakers of other languages\(^{[36]}\). Several important limitations of the current study must also be considered. Due to the availability of the clinical and normative data, we used only semantic verbal fluency tasks here, but did not include phonemic fluency, limiting the interpretation of our results. Further studies using a phonemic fluency test would extend our findings.

In summary, the current results revealed that the semantic VFT is sensitive in the differential diagnosis of various cognitive impairments, and is useful in the detection of MCI and VCIND.

**ACKNOWLEDGEMENTS**

This work was supported by the Science and Technology Commission of Shanghai Municipality (09DZ1950400, 08411951000) and the Fudan Elite Foundation.

Received date: 2012-05-07; Accepted date: 2012-08-01

**REFERENCES**


