Is semantic verbal fluency impairment explained by executive function deficits in schizophrenia?

Arthur A. Berberian,1,2 Giovanna V. Moraes,3 Ary Gadelha,1 Elisa Brietzke,1 Ana O. Fonseca,1 Bruno S. Scarpati,1 Marcella O. Vicente,1 Alessandra G. Seabra,4 Rodrigo A. Bressan,1 Acioly L. Lacerda1

1Programa de Esquizofrenia (PROESQ) and Laboratório Interdisciplinar de Neurociências Clínicas (LINC), Departamento de Psiquiatria, Universidade Federal de São Paulo (UNIFESP), São Paulo, SP, Brazil. 2Departamento de Psicologia Educacional, Centro Universitário FIEO (UNIFIEO), Osasco, SP, Brazil. 3Department of Cognitive & Linguistic Sciences, Brown University, Providence, RI, USA. 4Programa de Pós-Graduação em Distúrbios do Desenvolvimento, Universidade Mackenzie, São Paulo, SP, Brazil.

Objective: To investigate if verbal fluency impairment in schizophrenia reflects executive function deficits or results from degraded semantic store or inefficient search and retrieval strategies.

Method: Two groups were compared: 141 individuals with schizophrenia and 119 healthy age and education-matched controls. Both groups performed semantic and phonetic verbal fluency tasks. Performance was evaluated using three scores, based on 1) number of words generated; 2) number of clustered/related words; and 3) switching score. A fourth performance score based on the number of clusters was also measured.

Results: Individuals with schizophrenia produced fewer words than controls. After controlling for the total number of words produced, a difference was observed between the groups in the number of cluster-related words generated in the semantic task. In both groups, the number of words generated in the semantic task was higher than that generated in the phonemic task, although a significant group vs. fluency type interaction showed that subjects with schizophrenia had disproportionate semantic fluency impairment. Working memory was positively associated with increased production of words within clusters and inversely correlated with switching.

Conclusion: Semantic fluency impairment may be attributed to an inability (resulting from reduced cognitive control) to distinguish target signal from competing noise and to maintain cues for production of memory probes.

Keywords: Psychosis; schizophrenia; cognitive neuroscience; chronic psychiatric illness; language

Introduction

Verbal fluency is a cognitive function that allows retrieval of information from memory1 and can be measured by semantic (category) and phonological (letter) fluency tasks. While semantic fluency depends mainly on successful activation flow through the semantic network, phonemic fluency depends on the search and retrieval from the lexicon using phonemic or orthographic cues.2,3 In patients with schizophrenia (SZ), poor word fluency performance correlates with lower functioning levels,4 and may be the cause of formal thought disorder.5

One way of assessing verbal fluency is to measure two cognitive strategies commonly used by examinees, namely clustering (producing words within clusters or categories) and switching (ability to successfully switch between these clusters or categories).6-8 In patients with SZ, only the switching component of phonemic fluency has been found to be impaired8; however, assessment of semantic fluency revealed that both switching and clustering were affected8 as compared to healthy controls (HC). Nonetheless, when the total word output was considered, patients and HC appeared to employ the same cognitive strategies.6,10 Further examination of the differences in strategy implementation may help elucidate the cognitive processes involved in each fluency task.

Bozikas et al.6 have reported that a differential deficit in clustering, but not in switching, led to disproportionate worsening of semantic fluency in SZ. Authors have interpreted this as evidence of structural semantic knowledge impairment in SZ. However, Joyce et al.11 and Elvavag et al.10 have suggested that patients with SZ have executive deficits rather than semantic knowledge deficits. Both studies describe an improvement in semantic verbal fluency in SZ when executive demands are reduced. Therefore, it is still debatable whether verbal fluency deficits in SZ are caused predominantly by the structure of semantic knowledge12-14 or by inefficient search and retrieval strategies.10,15,16

Semantic categorizations are used more frequently because they are based on meaning. Bokat et al.15 argue that these categorizations should facilitate fluency as signals become more automatic. However, this automaticity may lead to increased noise during the semantic
therefore, to more frequent switching between retrieval cues to make up a memory probe. Consequently, working search is a dynamic process that relies on specific retrieval between working memory and switching.

of words related to clusters; 4) and negative correlation with SZ to that of age, sex, and education-matched controls.

switching effectiveness; 2) disproportionate impairments in a verbal fluency task. The following was investigated: 1) differences between groups in terms of clustering and switching effectiveness; 2) disproportionate impairments across semantic and phonetic fluency tasks within groups; 3) positive correlation between working memory and number of words related to clusters; 4) and negative correlation between working memory and switching.

Methods

This study was approved by the Research Ethics Committee of Universidade Federal de São Paulo (protocol no. 2155/08). After signing an informed consent form, subjects were assessed individually in a quiet room.

Participants and clinical assessment

The study included 260 participants: 141 individuals with SZ according to DSM-IV criteria (58.8% men) from an outpatient unit located at the Universidade Federal de São Paulo (UNIFESP), and 119 healthy age and education-matched controls (41.2% men) with no personal or familial (first-degree) history of psychiatric disorders. Diagnosis was confirmed using the Brazilian Portuguese version of the Semi-Structured Clinical Interview for DSM-IV - Axis I Disorders (SCID). Clinical assessment also encompassed application of the Positive and Negative Syndrome Scale (PANSS) and Global Assessment of Functioning (GAF).

All patients had been taking atypical antipsychotics at stable doses for at least 6 weeks prior to inclusion. Clinical and demographic characteristics of the sample are shown in Table 1.

Instruments and procedures

Semantic and phonemic fluency tests were administered.

In both, the subject had 60 seconds to produce as many words as possible within a specified category: animals in the semantic test and words starting with F, A, and S (each word beginning with a specific letter was shown independently for 60 seconds) in the fluency test. Participants received no guidelines on how to conduct the search and/or word production to ensure the measurement of spontaneous fluency.

Performance was evaluated using three scores, based on: 1) number of words generated (WG); 2) number of clustered/related words (RW); and 3) switching score (WG - RW + number of clusters). A fourth performance score based on the number of clusters was also measured (X). Repeated words were only counted once and errors were excluded.

Clusters were defined as at least three consecutive words belonging to the same subcategory: 1) number of words as possible within a specified category: animals in the semantic test and words starting with F, A, and S (each word beginning with a specific letter was shown independently for 60 seconds) in the fluency test.

Semantic and phonemic fluency tests were administered.

In both, the subject had 60 seconds to produce as many words as possible within a specified category: animals in the semantic test and words starting with F, A, and S (each word beginning with a specific letter was shown independently for 60 seconds) in the fluency test. Participants received no guidelines on how to conduct the search and/or word production to ensure the measurement of spontaneous fluency.

Performance was evaluated using three scores, based on: 1) number of words generated (WG); 2) number of clustered/related words (RW); and 3) switching score (WG - RW + number of clusters). A fourth performance score based on the number of clusters was also measured (X). Repeated words were only counted once and errors were excluded.

Clusters were defined as at least three consecutive words belonging to the same subcategory: a list of three or more farm animals was considered a semantic cluster and a list of three or more words that sounded alike or had only one different letter was defined as a phonetic cluster. All words representing a single subgroup were considered as RW. Switching (WG - RW + number of clusters) represented the number of switches made between clusters, and included single words. For example, the following sequence has three subcategories: chicken, goose, duck, shark, whale, crab, oyster, dog, lion, parrot, hawk, and gull. The first three words belong to the birds subcategory, the next four words belong to the category of fish and marine animals, and the last three to birds again. The words dog and lion do not make up a subcategory. This example thus shows 12 words, three semantic subcategories, five switches, 10 RW and a switching score of seven (WG - RW + number of clusters; 12 - 10 + 5 = 7).

A keep track task adapted from Yntema was used to evaluate the updating of working memory contents. In this study, the task consisted of showing several target categories (animals, colors, countries, distances, metals, and relatives) to the respondent on a computer screen. Fifteen words were then presented verbally and randomly, one at a time, during 1,500 milliseconds. Target categories remained on the computer screen. Each list included 2 or 3 exemplars from each of the six possible categories. Participants had to remember the last word presented that belonged to the remaining target categories (the first three
trials included four categories and the last three trials included five categories). For instance, if the target categories were colors, relatives, and animals, then, at the end of the trial, participants were expected to recall the last color, relative, and animal that had been presented to them. To be successful, respondents had to monitor the words and update their working memory representations. Before beginning the task, participants were allowed to look at all six categories and category exemplars to ensure that they knew to which category each word belonged and to practice twice with three and four target categories. The dependent measure was the ratio of words recalled correctly.

Previous studies have shown that this task is a valid option when comparing the cognitive performance of subjects with SZ to that of healthy individuals.25-27 Letter tasks were performed first, followed by semantic tasks. Finally, participants were assigned an updating task.

**Statistical analyses**

All statistical analyses were conducted using the SPSS version 15.0. A mean score was calculated for the phonemic performance. Normality of distribution was tested using the Kolmogorov-Smirnov test. Group comparisons were conducted with one-way analyses of variance (ANOVA) for continuous variables, and with chi-square test for categorical variables. We also conducted an analysis of covariance (ANCOVA) for the number of RW and the switching score, where the total number of words produced during both tasks was defined as a covariate.

Repeated measures ANOVA was conducted, having “group” as the between-group factor and “fluency type” (semantic, phonemic) as the within-subject factor for the total number of words. Repeated measures ANOVA was also conducted for the number of RW and the switching score, where the total number of words produced in both tasks was defined as a covariate. Because phonemic verbal fluency depends mainly on the search and retrieval processes in the lexicon using phonemic or orthographic cues instead of semantic networks, we conducted a single partial correlation analysis using only semantic switching and the number of RW to the keep track of task clusters, controlling for the total number of words per each group separately. For all statistical analyses, significance was set at p < 0.05, two-tailed.

**Results**

Based on the well-documented evidence showing better verbal processing for females,30 we compared gender performance between groups. No significant difference between males and females was identified in either the SZ group (letter fluency \[ t = 1.47; p = 0.14 \] and semantic fluency \[ t = 1.44; p = 0.15 \] respectively) or the control group (letter fluency \[ t = 1.72; p = 0.10 \] and semantic fluency \[ t = 1.13; p = 0.26 \] respectively).

Table 2 describes the mean WG, RW, and switching score obtained in the phonemic and semantic verbal fluency tasks for patients with SZ and HC. Between-group analysis showed that patients with SZ generated fewer total words than HC in both the semantic \( \left( F_{29,521} = 1.256, p < 0.001 \right) \) and phonemic \( \left( F_{13,043} = 1.256, p < 0.001 \right) \) tasks. The SZ group was also less successful at implementing clustering (semantic: \( F_{12,84} = 1.256, p < 0.001 \); phonemic: \( F_{3,73} = 1.256, p < 0.05 \)) and switching (semantic: \( F_{12,031} = 1.256, p < 0.001 \); phonemic: \( F_{7,080} = 1.256, p < 0.009 \)) cognitive strategies. However, these differences in strategy employment across groups disappeared when the total number of WG was considered (ANCOVA), except for the RW score achieved in the semantic task \( F_{1,258} = 6.31, p < 0.02 \).

Repeated measures ANOVA was used to predict the main effect of group \( F_{1,258} = 30.83, p < 0.001 \) and type of fluency \( F_{1,258} = 311.37, p < 0.001 \), and of the interaction of group vs. fluency type \( F_{1,258} = 5.87, p < 0.02 \) for total word production. SZ individuals produced significantly fewer words than controls and, overall, both groups generated significantly more words in the semantic fluency task than in the phonemic fluency task \( F_{311,37} = 1.258, p < 0.001 \). The significant interaction between group and fluency type suggests a disproportionate
impairment of semantic fluency compared to phonemic fluency in SZ.

There was a major effect of group (F\(_{1,258}\) = 25.80, p < 0.001) and fluency type (F\(_{1,258}\) = 378.072, p < 0.001) on the number of RW (cluster size), with patients and controls generating fewer RW during the phonetic task compared to the semantic verbal fluency. There was also a significant group vs. fluency type interaction (F\(_{1,258}\) = 14.74, p < 0.001), indicating that cluster size performance was worse in SZ.

In terms of the number of switches, there was a major effect of group (F\(_{1,258}\) = 14.59, p < 0.001) and fluency type (F\(_{1,258}\) = 15.47, p < 0.001), but not of group vs. fluency type interaction (F\(_{1,258}\) = 1.47, p = 0.226). Although SZ individuals created fewer clusters than controls, there was no significant interaction of group vs. fluency type for switching.

When we used the combined number of words produced in the two fluency tasks as a covariate for RW and number of switches, the significance of group vs. fluency type interaction between the number of RW (F\(_{1,258}\) = 0.32, p = 0.573) and number of switches (F\(_{1,256}\) = 0.41, p = 0.52) disappeared.

A weak and positive correlation between number of RW and the keep track task was found (r = 0.20; p < 0.03) in the SZ group. No association between these two measures (r = 0.06; p = 0.32) was found in the HC group. A weak and marginal negative correlation between semantic switching and the keep track task was found (r = -0.15; p = 0.08) in the SZ group. No association between these two measures (r = -0.07; p = 0.46) was found in the HC group.

Since SZ individuals performed worse than controls at the keep track task (Table 2), we used this updating measure as a covariate of RW and number of switches. Group vs. fluency type interaction was not associated with either number of RW (F\(_{1,257}\) = 0.008, p = 0.87) or number of switches (F\(_{1,257}\) = 0.000, p = 0.98).

**Discussion**

As expected, our data indicated a verbal fluency deficit in SZ. SZ individuals generated fewer total words than HC in the semantic and phonemic tasks. However, after correcting for this reduced output, we observed that patients employed the same cognitive strategies as controls (number of clusters and switching score), although a difference remained in the number of exemplars per cluster in the semantic test. These results may reflect the pronounced worsening of semantic fluency in the SZ group. We also obtained evidence – even though based on weak associations – of involvement of the working memory process, since the number of RW was positively correlated with the keep track task, whereas switching correlated only marginally and inversely with the keep track task in the SZ group. After correcting for working memory, SZ individuals engaged in cognitive strategies (number of clusters and switching) to the same extent as controls.

Similar data have been reported in the literature and many studies see these reports as evidence of semantic impairment in SZ. Semantic categorizations are used more frequently because they are based on meaning. Bokat & Goldberg suggest that this makes the semantic task easier than the phonetic task because categorizations are rehearsed more often and require less controlled processing. Rehearsal makes for stronger, more automated associations, thereby promoting fluency. This is a useful assumption considering that subjects reportedly provide more words in semantic tests. However, it is possible that the semantic network is also more widespread, and may need controlled processing to provide adaptive responses. Since both groups generated the same number of semantic fields, a fact that corroborates, at least to a certain extent, that the semantic storage has been preserved as shown by Troyer et al. and Rohrer et al., the reduced number of exemplars per cluster in the semantic milieu may be the result of a broader semantic network activation range. Whereas certain semantic categorizations facilitate access to other categorizations, they may also increase noise within the system, demanding more cognitive control for adaptive behavior.

With this in mind, it is possible to speculate that both groups generated more words in the semantic fluency task because of the broader activation range (there is more to choose from), and that SZ patients were proportionally worse at semantic clustering due to their inability to distinguish between signals (more competition). Furthermore, this would explain why SZ performance improves with cuing, task repetition, or prior organization strategies. More specifically, these tactics reduce “executive” demands, enabling

**Table 2** Descriptive analysis of variations in verbal fluency tasks and the keep track task for schizophrenia (n=141) and healthy groups (n=119) after ANOVA correction

<table>
<thead>
<tr>
<th>Variable</th>
<th>SZ</th>
<th>HC</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>F (df=1.258)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Letter fluency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total words</td>
<td>9.64±3.84</td>
<td>11.31±3.84</td>
<td>0.27</td>
<td>0.33</td>
<td>13.043</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Related words</td>
<td>2.07±1.88</td>
<td>2.51±1.76</td>
<td>0.78</td>
<td>0.27</td>
<td>3.730</td>
<td>0.05</td>
</tr>
<tr>
<td>Number of clusters</td>
<td>6.22±5.65</td>
<td>7.54±5.30</td>
<td>0.45</td>
<td>-0.57</td>
<td>7.070</td>
<td>0.008</td>
</tr>
<tr>
<td>Switching</td>
<td>8.16±3.33</td>
<td>9.28±3.39</td>
<td>0.33</td>
<td>-0.14</td>
<td>7.080</td>
<td>0.008</td>
</tr>
<tr>
<td><strong>Semantic fluency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total words</td>
<td>14.06±4.61</td>
<td>17.14±4.47</td>
<td>0.07</td>
<td>0.41</td>
<td>29.521</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Related words</td>
<td>6.62±4.20</td>
<td>9.30±4.72</td>
<td>0.27</td>
<td>-0.26</td>
<td>12.843</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Number of clusters</td>
<td>2.89±1.12</td>
<td>3.39±1.13</td>
<td>0.02</td>
<td>-0.78</td>
<td>23.40</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Switching</td>
<td>8.91±3.67</td>
<td>10.70±4.61</td>
<td>0.22</td>
<td>0.00</td>
<td>12.031</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td><strong>Keep track task</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13.16±3.82</td>
<td>16.81±3.16</td>
<td>0.32</td>
<td>0.26</td>
<td>18.79</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Data presented as mean ± standard deviation, unless otherwise specified. df = degrees of freedom; HC = healthy controls; SZ = schizophrenia.
optimum interpretation of concurrent representations. This suggests that certain cognitive deficits in SZ might reflect a highly disorganized and inconsistent network that is unable to distinguish target signal from competing noise.

In spite of having used a plethora of different tasks, several authors have become strong advocates of this idea. For example, an fMRI study by Tregellas et al.\textsuperscript{35} found that SZ patients were unable to appropriately engage the task-relevant network to the same degree as controls, and that this difference was further intensified in the presence of increased distracting noise. Furthermore, the authors attributed this to a basal and hippocampal hyper-responsive-ness to stimuli causing incongruous cortical recruitment. Similarly, EEG studies by authors such as Suazo et al.\textsuperscript{36} reported increased gamma band activation and decreased theta band activation during a verbal memory task. Theta band activation has been proposed as an important top-down influence and might play a major role in adjusting SNRs and coordinating the neural circuits involved in higher cerebral functions. Curiously, this same pattern is not found in bipolar patients.\textsuperscript{35} Winterer & Weinberger\textsuperscript{17} reviewed this pattern and identified increased response variability and unfocused or unstable response circuits in SZ patients compared to controls. In other words, abnormal excitation in SZ increases spontaneous discharge (noise), thereby decreasing network specificity (SNR). The authors emphasize the role of dopamine, glutamate, and GABA in maintaining optimal SNR, suggesting that deficient levels of these neurotransmitters could lead to the deficits seen in SZ. The combination of these deficits suggests that subspherical hyper-activation and concurrent cortical hypo-activation are possible underlying causes of cognitive deficit in SZ.

According to the cue-maintenance hypotheses,\textsuperscript{18} the ability to maintain focus on one cue while ignoring other potentially distracting ones\textsuperscript{19} is crucial. Our results show that the number of RW was weakly associated with working memory. Additionally, a trend towards increased switching was seen at lower levels of keep track tasks. Although these were weak associations, the first correlation suggests that individuals who are better at cue-maintenance are able to produce more words within clusters; this is probably so because of their focus, which allows these individuals to enhance their memory probe. The second correlation suggests that a compensatory search mechanism is set in motion, and that switching may be associated with loss of cue-maintenance. This hypothesis is reinforced by equivalent cognitive strategy (number of clusters and switching) implementation between groups after correcting for working memory differences, thereby implicating declined cognitive control for the witnessed verbal fluency impairments in SZ.

The main limitation of this study has to do with the method used to score switching. We used a method based on the Troyer et al.\textsuperscript{7} categorization, which was replicated by Robert et al.\textsuperscript{8} According to that approach, categories are hand-coded, and a detailed list of proposed categories is likely to be used to access semantic memory (e.g., fishes and marine animals). However, this method may not be accurate enough to assess additional switching behavior because it does not specify how search is carried out within a category. Indeed, Hills et al.\textsuperscript{18} have proposed that the memory probe is dynamic and can change over the course of the retrieval period, swinging between a currently accessible cognitive representation area (local search) and new areas (global search). The first cue to be established is the global retrieval cue, which defines the boundaries of the search space (e.g., animals). Then, as the item is recovered (e.g., “dog”) the local cue is defined and the word increases the activation of other items that are semantically closer to the most recent cue (e.g., “cat”). When an individual fails to retrieve an item, his or her memory probe eventually loses the item’s local cue, returning to its global form until a new local cue is received. The use of this scoring method could provide the sensitivity required to analyze change over the course of the retrieval period and to provide new insights into how SZ affects the exploration mechanism of memory search.

Another limitation of this study is that only one trial was used to evaluate semantic verbal fluency (animals). We chose this procedure because we believed that additional semantic trials would increase the risk of eliminating or reducing the differences detected. However, considering that similar results have been extensively reported in the literature,\textsuperscript{6,11,13,16} we feel that our results are reliable. Another limitation is that our study did not examine strategy implementation according to clustering and switching subtypes in more detail. Differently from Abwender et al.,\textsuperscript{1} we only scored task-consistent clustering, i.e., phonemic clustering in phonemic fluency tasks, but did not score task-discrepant clustering, i.e., semantic clustering in phonemic fluency tasks. A more detailed analysis of cognitive strategies could have revealed deficits that are beyond the scope of this study. However, due to the reduced output seen in SZ, it would have been inefficient to distinguish and further investigate these strategy subtypes. Moreover, because our methodology was not designed to examine executive function abilities directly, it would have been impossible to determine whether the semantic clustering deficit reported in this study could be related to weaker executive abilities.

In summary, compared to HC, individuals with SZ accessed the same number of semantic fields, but with smaller cluster sizes. The disproportionate impairment in semantic fluency found in SZ may have been caused by insufficient access and retrieval capabilities (due to the difficulty in selecting words for clustering) and by inhibiting distractors (resulting from increased SNR)\textsuperscript{36} rather than by reduced semantic function in SZ per se\textsuperscript{11,15,33} The weak correlation between the number of RW and the working memory-associated trend to increase switching is consistent with working memory models and represents a measure of the inability to focus and/or inhibit distracting information. Instead of being a reflection of insufficient semantic store or absence of search and retrieval strategies, these deficits are the consequence of the ways in which the impairment of the executive function manifests itself. This idea may help clarify the cognitive processes involved in SZ and identify targets for cognitive enhancement strategies.

Acknowledgements

We would like to thank Fundação de Amparo e Pesquisa do Estado de São Paulo (FAPESP; grants 2011/50740-5 and 2007/58630-9) for its financial support. Special thanks
Disclosure
The authors report no conflicts of interest.

References