Journal of the International Neuropsychological Society (2003), **9**, 384–393.

DOI: 10.1017/S1355617703930049

Category fluency in first-episode schizophrenia

TANIA GIOVANNETTI,[1] RITA Z. GOLDSTEIN,[2] MATTHEW SCHULLERY,[3] WILLIAM B. BARR,[4] AND ROBERT M. BILDER[5]

[1]Moss Rehabilitation Research Institute, Philadelphia, Pennsylvania
[2]Brookhaven National Laboratories, Upton, New York
[3]Widner University, Chester, Pennsylvania
[4]NYU Comprehensive Epilepsy Center, New York, New York
[5]Research Department, Hillside Hospital division of the North Shore-Long Island Jewish Health System, Glen Oaks, New York

(Received October 16, 2001; Revised April 22, 2002; Accepted April 23, 2002)

Abstract

Animal word list generation (ANWLG) was administered to 47 first-episode schizophrenia (FES) participants and 31 controls. Fifty-nine left temporal lobe epilepsy (LTLE) participants were included as a comparison group with known temporal lobe damage and expected semantic deficits. Semantic knowledge was assessed with the Association Index (AI), a measure of the semantic relatedness of *all* consecutive ANWLG responses. Neuropsychological tests of language and executive functioning were also administered. Results showed that both FES and LTLE groups generated fewer ANWLG responses than controls, but only the LTLE participants obtained a lower AI relative to controls. FES participants did not differ from controls on the AI. FES and LTLE groups produced fewer semantic subcategories (clusters), however, only the LTLE group produced fewer words per subcategory compared to controls (cluster size). FES participants produced a higher rate of perseverative responses compared to the other groups. Finally, correlation analyses showed that for FES participants both executive and language tests significantly correlated with ANWLG total responses, while the correlation between ANWLG and only 1 language test was significant for LTLE participants. Taken together, the results suggest that reduced ANWLG output in FES participants may be best conceptualized as a deficit in the executive component of word list generation (i.e., semantic search/access, response monitoring) or global cognitive impairment. (*JINS*, 2003, *9*, 384–393.)

Keywords: Schizophrenia, Semantic knowledge, Animal word list generation, Category fluency, First episode schizophrenia

INTRODUCTION

Recently, investigators have attempted to link the hallmark features of schizophrenia, namely disorganized thinking and loosening of associations (American Psychiatric Association, 1994), to semantic knowledge deficits (Goldberg et al., 1998; Rossell et al., 1999). Studies cite poor performance on animal word list generation (ANWLG) and indices derived from this task as evidence for semantic knowledge deficits in schizophrenia (Aloia et al., 1996; Goldberg et al., 1998; Gourovitch et al., 1996; McKay et al., 1996; Paulsen et al., 1996; Rossell et al., 1999). However, these studies have not considered several methodological issues that may have influenced their results, including the dependent variables used to assess the semantic component of ANWLG performance, inclusion of only chronic patient samples, and the lack of comparison groups with semantic processing deficits. Each of these potential methodological limitations is reviewed below and addressed in the design of the present study.

Reprint requests to: Tania Giovannetti, Ph.D., Moss Rehabilitation Research Institute, Korman Building, Suite 203B, 1200 West Tabor Road, Philadelphia, PA 19141. E-mail: Giovannt@einstein.edu

ANWLG Dependent Variables

Several authors have interpreted impaired ANWLG in schizophrenia not as a deficit in semantic knowledge, but as an impairment of willed action (Frith, 1992) or executive search processes (Allen et al., 1993; Joyce et al., 1996;

Zakzanis et al., 2000). Interpretation of reduced output on ANWLG is ambiguous due to the multifactorial nature of this task. Gruenewald and Lockhead (1980) proposed a two-component model of category fluency performance. Their model is based on the observation that healthy, control participants initially produce a cluster of highly semantically related responses (i.e., responses that share many similar attributes) in quick succession. Then, after a comparatively long response latency, participants switch into a new cluster of highly related responses. Gruenewald and Lockhead (1980) concluded that the ability to generate clusters of semantically meaningful responses (semantic knowledge) is independent from the ability to search and shift from one cluster to another (semantic search/access). Thus, reduced ANWLG output may be secondary to degradation of semantic knowledge or impaired search/access (see also Warrington & Shallice, 1979). Moreover, the search/access and semantic knowledge components have been associated with different brain regions, with the prefrontal cortex associated with the former and temporal lobe linked to the latter (Troyer et al., 1997, 1998a, 1998b).

Several investigators have acknowledged the multifactorial nature of category fluency and have turned to multidimensional scaling and additive tree clustering techniques to measure the integrity of semantic knowledge independent from search/access processes (Aloia et al., 1996; Chan et al., 1993; Paulsen et al., 1996; Rossell et al., 1999). In brief, these statistical techniques assess the order in which a subset of 11 to 17 of the most common ANWLG responses is produced (i.e., *cat*, *dog*, *cow*, *horse*, etc.). Based on the notion of automatic spreading of activation, it is assumed that animals that are semantically related (e.g., *horse* and *cow*) are produced in close proximity (i.e., with fewer intervening words) relative to animals that are not so strongly associated (e.g., *horse* and *cat*). In general, results have shown that schizophrenia participants demonstrate a less stable (two-dimensional) organization of this subset of responses and fail to group the responses into meaningful clusters relative to healthy controls. These data have been interpreted as evidence of impaired semantic knowledge in schizophrenia (Aloia et al., 1996; Paulsen et al., 1996; Rossell et al., 1999).

Multidimensional scaling and related techniques offer a rigorous method to quantify the semantic relatedness of ANWLG responses. One limitation of these techniques, however, is that only a subset of the most frequent responses is analyzed. For example, Aloia et al. (1996) reported that schizophrenia patients generated an average of 14 responses of which an average of 5.7 (or 41%) were included in the multidimensional scaling analysis. A second limitation is that patients consistently generate fewer ANWLG responses relative to controls. Thus, patients generally generate fewer of the analyzed target responses (Aloia et al., 1996; Paulsen et al., 1996; Rossell et al., 1999). Although some have argued that the statistical methods used in multidimensional scaling are robust irrespective of group differences in the number of target words analyzed (Chan et al., 1993), others have acknowledged the potential bias in these methods (Rossell et al., 1999).

Several studies have explored the semantic integrity of ANWLG output with methods that include all responses (Moelter et al., 2001; Robert et al., 1998; Zakzanis et al., 2000). For example, Robert et al. (1998) and Zakzanis et al. (2000) recorded the number of consecutive responses generated from the same semantic subcategory or cluster (e.g., *farm animals*, *birds*, etc.) and the number of times participants switched from one semantic subcategory to another.[1] Both studies reported that schizophrenia patients produced fewer words within a semantic subcategory cluster and switched from one subcategory to another less often relative to healthy controls. Although Robert et al. (1998) interpreted their findings as evidence for both impaired semantic knowledge and retrieval, Zakzanis et al. (2000) reported that the impairment in switching was greater than that of cluster size and suggested a more significant deficit in search/access (i.e., frontal lobe functions) relative to semantic knowledge (i.e., temporal lobe function). Finally, Moelter et al. (2001) included both multidimensional scaling and a measure that quantifies the number of semantic attributes shared by consecutive responses (i.e., Association Index; Giovannetti et al., 1997) in their study of ANWLG. Multidimensional scaling results were consistent with previous reports, and the Association Index showed that schizophrenia patients' ANWLG responses shared significantly fewer attributes compared to controls' responses. They concluded that both search/access and semantic knowledge are impaired in schizophrenia.

Disease Chronicity

Of the studies that have examined all ANWLG responses, only chronic patient samples were included (Moelter et al., 2001; Robert et al., 1998; Zakzanis et al., 2000). We believe the issue of disease chronicity is potentially relevant and warrants investigation. For example, Paulsen et al. (1996) reported that patients with late-onset schizophrenia (i.e., onset after age 45) showed relatively preserved semantic organization of ANWLG output relative to patients with an earlier onset of the disease. The late and early onset patients were the same age at the time of testing and generated an equal number of words on ANWLG. They differed only in medication regimen and negative symptoms (i.e., more early onset patients received anticholinergics and obtained higher negative symptom ratings). The present study explores the integrity of semantic knowledge in first episode schizophrenia patients to assess whether evidence of semantic deficits in chronic schizophrenia patients reflect impairment secondary to schizophrenia *per se* or from the long-term, global cognitive effects of medication, and/or institutionalization (see Friedman et al., 1999).

[1]Troyer (2000) and her colleagues (1997a, 1998a, 1998b) have extensively researched this scoring method with a wide range of populations, including dementia, normal aging, and focal cortical lesions.

Semantic Knowledge-Impaired Comparison Group

Lastly, studies of ANWLG in schizophrenia have not included comparison groups of patients expected to show semantic knowledge impairments (i.e., Alzheimer's disease, left temporal lobe epilepsy). A semantically impaired comparison group would provide a standard by which to evaluate schizophrenia patients' performance on ANWLG semantic indices, while controlling for reduced fluency and the effects of general cognitive impairment.

Patients with left temporal lobe epilepsy (LTLE) are an ideal comparison group for several reasons. First, the temporal lobes, which are affected in LTLE, are considered to be the site of semantic storage (Chertkow et al., 1997; Garrard et al., 1997; Hart & Gordon, 1990; Papps et al., 2000; Troyer et al., 1997, 1998a, 1998b), and investigators have demonstrated impaired semantic knowledge in LTLE patients on a range of tasks (Bell et al., 2001; Helmstaedter et al., 1997), including category fluency (Tröster et al., 1995). For example, Tröster et al. (1995) have shown that LTLE participants demonstrate smaller clusters and more superordinate category labels (e.g., *fruit vs. apple*) relative to controls on the Supermarket Fluency Test. Lastly, LTLE participants tend to be considerably younger than other patient populations with semantic deficits (i.e., Alzheimer's disease), making them a more appropriate comparison group for FES participants (see also Gold et al., 1994; Seidman et al., 1998).

Present Study

To determine whether reports of semantic knowledge deficits in previous studies were due to the fact that only a portion of output was analyzed or that only chronic patients were assessed, the present study examines semantic knowledge by analyzing all ANWLG responses produced by first episode schizophrenia participants. Moreover, in addition to healthy controls, participants with left temporal epilepsy (LTLE) were studied as a semantically impaired comparison group. Performance of the comparison group will demonstrate how participants with semantic knowledge deficits from temporal lobe dysfunction will perform on our ANWLG indices. If both FES and LTLE participants are impaired on our ANWLG indices of semantic knowledge, then we will show support for semantic knowledge degradation in schizophrenia. If, however, only LTLE participants are impaired on the semantic component of ANWLG, then the results will suggest that semantic knowledge is relatively preserved in FES, and that the methodological issues addressed above might explain the previous reports of semantic knowledge degradation in schizophrenia.

Finally, we also explored the relationship between background tests of executive functions and language and ANWLG performance. Tests of executive functioning and language are associated with the prefrontal cortex and temporal lobe respectively. Therefore, correlations between these tests and ANWLG performance will provide additional evidence for the neuropsychological functions that contribute to reduced output for each group. If ANWLG is reduced due to semantic knowledge/temporal lobe deficits, then ANWLG should significantly correlate with language tests. By contrast, if output is low due to search/access/frontal lobe deficits, then tests of executive functioning and ANWLG should significantly correlate.

METHODS

Research Participants

Forty-seven first episode schizophrenia (FES) participants were recruited from Hillside Hospital after admission for first episode of psychotic illness. Twenty-seven participants were men (57%) and 32 were right-hand dominant (68%). Participants had less than 12 weeks cumulative neuroleptic treatment. All met Research Diagnostic Criteria for Schizophrenia (Spitzer et al., 1978; subtypes included 37 paranoid, 3 disorganized, and 7 undifferentiated) and all were tested after initial stabilization of psychosis. These participants were enrolled in the Prospective Study of Psychobiology in First Episode Schizophrenia (previously reported in Bilder et al., 1991, 1992, 1995, 2000; Goldstein et al., 2002; Lieberman et al., 1992, 1993a, 1993b; Robinson et al., 1999a, 1999b).

Fifty-nine left temporal lobe epilepsy (LTLE) patients were recruited from Long Island Jewish Medical Center Epilepsy Clinic. All participants were tested as part of a comprehensive evaluation for epilepsy surgery. Participants were considered to have medically intractable complex partial seizures of left temporal lobe origin. All participants were evaluated with continuous video-scalp electroencephalographic (EEG) monitoring and seizure focus was confirmed from the results of ictal and inter-ictal EEG recordings. All participants subsequently underwent a left temporal lobectomy. On average, participants had suffered from epilepsy for 14.2 years ($SD = 9.6$) and were prescribed various anticonvulsant medications. Twenty-four participants were men (41%) and all were right-hand dominant.

Thirty-one healthy controls were recruited from the Long Island Jewish Medical Center/Hillside Hospital community through announcements in local newspapers and within the medical center. They were free of mental disorders as determined by using the Schedule for Affective Disorders and Schizophrenia (SADS) Lifetime Version Interview, physical examination, and urinalysis (see Lieberman et al., 1992). Participants reported no history of substance abuse or neurologic/psychiatric/medical illness. Twenty-one of the controls were men (68%) and 23 were right-hand dominant (74%). Data from these participants have been reported in previous studies (Bilder et al., 1991, 1992, 1995, 2000; Lieberman et al., 1992).

Table 1. Demographic characterization of the groups

Characteristic	First episode schizophrenia ($N = 47$) *M*	(*SD*)	Left temporal lobe epilepsy ($N = 59$) *M*	(*SD*)	Controls ($N = 31$) *M*	(*SD*)
Age	25.76	6.77	33.1	9.17	25.2	6.07
Education	13.26	1.95	12.66	2.33	15.0	1.48
WAIS–R FSIQ	84.94	12.51	84.69	12.62	109.3	11.51

Demographic Characterization of the Groups

Table 1 shows that the LTLE group was significantly older than the FES and control groups, [$F(2{,}134) = 16.0$, $p < .001$]. The FES and LTLE groups had fewer years of education, [$F(2{,}134) = 12.93$, $p < .001$], and a lower Wechsler Adult Intelligence Scale–Revised (WAIS–R; Wechsler, 1987) Full Scale IQ [$F(2{,}134) = 46.04$, $p < .001$], compared to control participants. The distribution of men and women differed [$\chi^2(2) = 6.7$, $p = .036$]. All demographic variables were included as covariates in subsequent between group analyses.

Animal Word List Generation (ANWLG)

Participants produced as many different animal names as possible in 60 s. The following 5 dependent variables were obtained:

1. *Total Responses:* The total number of responses, including perseverations and extra-category intrusion responses (e.g., *chair*).
2. *Association Index (AI):* Subjects' responses were characterized on each of the following categories as described in Giovannetti et al. (1997): size (*big*, *small*); geographic location (*foreign*, *North America*); diet (*herbivore*, *carnivore*, *omnivore*); zoological class (*insect*, *mammal*, *bird*, etc.); habitat (*farm*, *Africa/jungle*, *widespread*, etc.); and biological order/related groupings (*feline*, *canine*, *bovine*, etc.).[2] The AI is calculated by summing the number of shared category attributes between all successive responses and then dividing by the total number of words generated minus one. Using this scoring technique, a string of highly associated responses (e.g., *tiger*, *jaguar*, *leopard*, etc.) share many attributes and yield a higher AI than a string of unrelated responses (e.g., *tiger*, *alligator*, *pigeon*). In sum, the AI is a measure of the strength of semantic association between all consecutive responses. Support for this index as a measure of semantic knowledge was obtained in a previous study that showed the AI, but not total ANWLG responses, was significantly lower in participants with Alzheimer's disease compared to participants with subcortical ischemic vascular dementia (see Giovannetti et al., 1997, for more details).
3. *Number of Clusters (Clusters):* A cluster is defined as a group of two or more consecutive responses that share *any* four or more attributes. The decision to define clusters on the basis of four attributes was made to ensure a minimum degree of association between successive responses. For example, in a previous study (Giovannetti et al., 1997) we observed that clusters defined on the basis of fewer than four attributes sometimes resulted in groupings that were not clearly meaningful (i.e., *frog* and *ant* share the three attributes *small*, *native*, and *widespread*). By contrast, upon inspection, animals that shared four or more attributes were much more obviously associated. Although admittedly arbitrary, this cluster score is significantly reduced in elderly dementia participants relative to healthy, age-matched controls (Giovannetti et al., 1997)

   The number of clusters generated was summed for each participant. Based on the conceptualization of Grueneewald and Lockhead (1980), this index measures the executive ability necessary to search and shift from one semantic subcategory to another (see Giovannetti et al., 1997).
4. *Cluster Size:* The number of words in a cluster divided by the number of clusters. This variable assesses the "automatic" activation of exemplars from a semantic subcategory once the subcategory is accessed.
5. *Perseverative Response Rate*[3]*:* The percent of ANWLG responses that were perseverations. We chose this measure instead of the total number of perseverative errors to control for the difference in ANWLG output between patient groups and controls. The perseverative rate was calculated by dividing the number of perseverations by ANWLG Total Responses (i.e., correct responses + perseverations + other errors), and then multiplying by 100.

Neuropsychological Assessment

In addition to the WAIS–R and ANWLG, tests of language and executive functioning were administered. The 60-item

[2] We have compiled a large electronic dictionary of animals that lists their relevant attributes to ensure consistency across raters scoring the AI.

[3] Extra-category intrusion errors were extremely rare across all groups; therefore, these errors were not analyzed.

version of the Boston Naming Test (BNT; Kaplan et al., 1983), Sentence Repetition (Benton & Hamsher, 1978; Spreen & Benton, 1969), and Token Test (Benton & Hamsher, 1978; DeRenzi & Falgoni, 1978) were administered to assess basic language skills associated with temporal lobe functions. The number of correct items was the dependent variable for the BNT (maximum = 60) and Sentence Repetition (maximum = 22). The dependent variable for the Token test was the number of correctly executed commands after the first or second administration (maximum = 78).

The Trail Making Test and Wisconsin Card Sorting Test (Grant & Berg, 1948) were administered as indices of executive control/prefrontal lobe functions. The dependent variables for the Trail Making Test included the time (in seconds) to complete Part A (Trails–A) and the time to complete Part B (Trails–B). For the WCST the dependent variables were the number of perseverative errors (WCST–perseverations) and the number of categories successfully completed (WCST–categories).

Data Analysis

Statistical analyses were performed with the SPSS package (Release 9.0.1). Variables that were not normally distributed were transformed for parametric statistical analysis: clusters (log), BNT (square root), Trails–A (square root), Trials–B (log). Between-group differences for all normal (and transformed) variables were analyzed with a series of univariate analyses of variance (ANOVAs). Scheffé tests were used for all *post-hoc* comparisons. Demographic variables that differed across the groups (i.e., age, education, and FSIQ) were included as covariates in all ANOVAs. Nonparametric Kruskal Wallis tests with *post-hoc* Mann-Whitney *U* tests were used with variables that were not normal and unable to be transformed (cluster size, perseverations, Token Test, WCST–Categories, WCST–perseverations). Pearson correlations were used to analyze relationships between ANWLG correct responses and neuropsychological test variables. Spearman rank-order correlations were used with nonnormal data. Bonferroni corrected *p* values were used to interpret significance when multiple comparisons/correlations were performed.

RESULTS

Total Responses

The mean Total Responses across the groups is shown in Table 2. Analyses revealed a significant between group difference for Total Responses [$F(2,134) = 42.6$, $p < .001$]. *Post-hoc* analyses showed that the FES and LTLE groups produced fewer correct responses compared to controls ($p <$ .001; effect sizes > 1.3 for both). This result remained significant when age, education and FSIQ were entered as covariates [$F(3,133) > 7.4$, $p < .001$ for all].

Table 2. Group means and standard deviations for total responses, association index, and clusters

	First episode schizophrenia ($N = 47$)		Left temporal lobe epilepsy ($N = 59$)		Controls ($N = 31$)	
Measure	*M*	*SD*	*M*	*SD*	*M*	*SD*
Correct responses	16.0	5.6	14.4	4.9	24.9	5.5
Association index	3.1	0.68	2.8	0.65	3.3	0.47
Clusters	3.7	1.6	3.4	1.5	6.2	2.1

Association Index (AI)

The mean AI across the groups is reported in Table 2. There was a significant effect of group for AI [$F(2,134) = 6.1$, $p = .003$]. *Post-hoc* tests showed that only the LTLE group obtained a lower AI compared to controls ($p = .005$; effect size = .79). There was a trend suggesting that FES participants obtained a higher AI relative to the LTLE participants ($p = .079$; effect size = .45), and there was no difference between the FES group and controls. This effect remained significant after demographic variables were entered as covariates [$F(3,133) > 4.20$, $p < .017$ for all].

Clusters

A significant between-groups difference was also observed for clusters [$F(2,134) = 23.1$, $p < .001$]. FES and LTLE participants produced fewer clusters relative to controls ($p <$ .001, effect sizes > 1.25 for both), but there was no difference between the LTLE and FES groups (see Table 2). This difference remained when demographic variables were included as covariates [$F(3,133) > 7.43$, $p < .001$ for all].[4]

Cluster Size

Figure 1 illustrates the distribution of cluster size by group. Analyses showed that the groups were significantly different on this variable [Kruskal-Wallis $\chi^2(2) = 8.35$, $p = .015$; mean ranks: FES = 69.93, LTLE = 59.89; controls = 85.08]. Controls generated significantly larger clusters relative to LTLE participants ($z = -2.9$, $p < .01$), but no other comparison was significant. The influence of demographic variables on the size of clusters was assessed with Spearman Rank Order correlations. These analyses showed cluster size was significantly positively correlated with education ($r =$ .23, $p = .005$), and Full Scale IQ ($r = .32$, $p < .001$). Therefore, it is possible that these variables may be mediating the difference in cluster size between the LTLE and control group.

[4]These findings were also obtained when the nontransformed cluster variable was analyzed with nonparametric Kruskal-Wallis and Mann-Whitney tests.

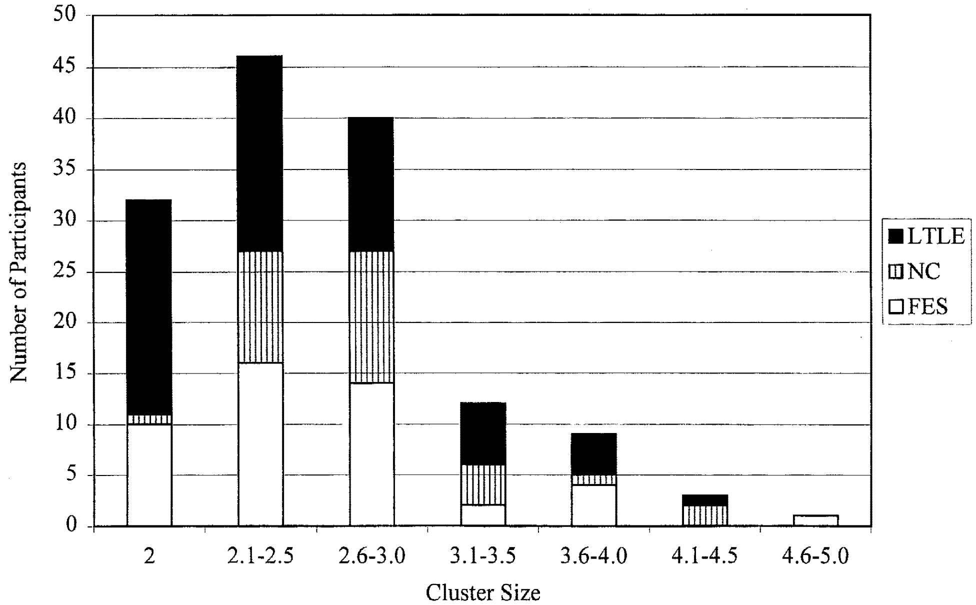

Fig. 1. Distribution of cluster size by group.

Perseverative Response Rate

Figure 2 shows the distribution of the perseverative response rate by group. The between group analysis was significant [Kruskal Wallis $\chi^2(2) = 7.8$, $p = .020$; mean ranks: FES = 79.97, LTLE = 63.19, controls = 63.42]. *Post-hoc* tests showed that the FES group made more perseverative responses than both LTLE ($z = -2.5$, $p = .013$) and NC ($z = -2.1$, $p = .033$) participants. There was no difference between the LTLE and control groups. The correlations between perseverative response rate and age, education, and FSIQ were not significant; thus the between group differences cannot be attributed to demographic variables.

Neuropsychological Assessment and Correlation Analyses

The group means (or mean ranks) and between group differences on the neuropsychological tests of language and executive functions are reported in Table 3. As can be seen the patient groups were impaired relative to controls on all

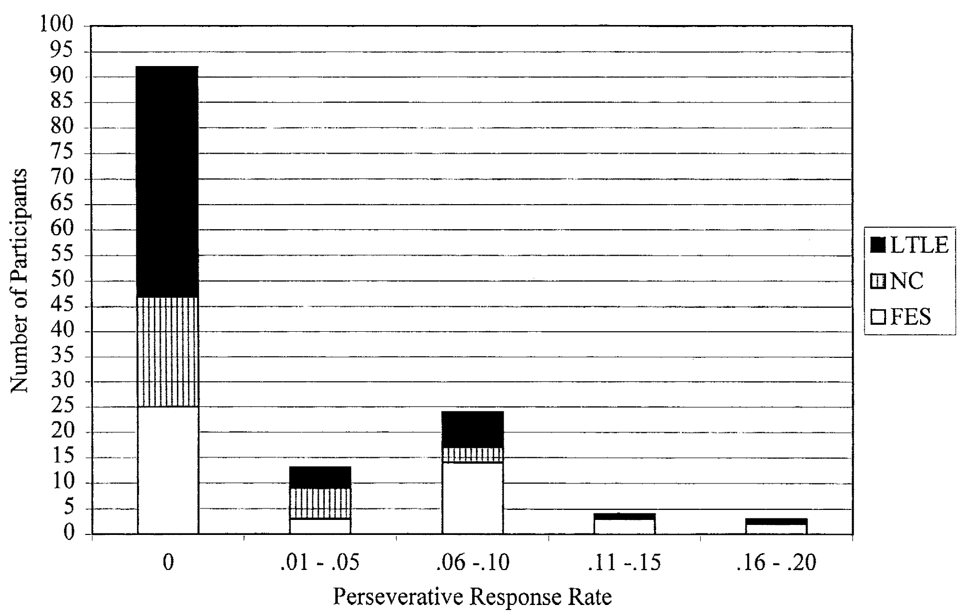

Fig. 2. Distribution of perseverative response rate by group.

Table 3. Means (or mean ranks), standard deviations, and between-group differences on background neuropsychological tests

Measure	First episode schizophrenia		Left temporal lobe epilepsy		Controls		Analyses	Interpretation
	M	*SD*	*M*	*SD*	*M*	*SD*		
Language								
Sentence Repetition	14.0	2.6	13.0	2.5	16.2	2.2	$F(2,124) = 15.9, p < .001$	LTLE = FES < NC
Token Test*	43.1		5.5		58.2		$\chi^2(2) = 37.7, p < .001$	LTLE < FES < NC
Boston Naming Test	45.7	9.2	38.9	10.6	54.7	3.7	$F(2,129) = 34.4, p < .001$	LTLE < FES < NC
Executive functions								
Trails–A	69.6	28.7	38.8	18.6	44.2	17.6	$F(2,131) = 28.9, p < .001$	FES > LTLE > NC
Trails–B	108.9	59.6	92.5	44.6	61.4	29.6	$F(2,129) = 14.4, p < .001$	FES = LTLE > NC
WCST–Categories*	51.5		69.9		83.2		$\chi^2(2) = 16.1, p < .001$	FES < LTLE = NC
WCST–Perseverations*	86.4		62.6		40.9		$\chi^2(2) = 27.5, p < .001$	FES < LTLE < NC

*Mean ranks obtained from Kruskal Wallis Test are reported for all variables that were analyzed with nonparametric analyses.

measures, with the exception of WCST–categories, on which LTLE participants performed comparably to controls. In general, LTLE participants performed significantly worse than FES participants on language tests, while the opposite was observed for tests of executive functions.

Table 4 shows the correlations between neuropsychological tests and total responses on ANWLG for each group. For the FES group, there was a significant correlation between total responses and both executive and language tests. For the LTLE group, only the correlation between total responses and a test of language was significant. Correlations between total responses and tests of executive functions did not reach significance after Bonferroni correction. Among controls, the correlation between Total Responses and Trails B just missed statistical significance.

DISCUSSION

FES and LTLE participants generated significantly fewer ANWLG correct responses relative to controls. However, unlike the LTLE group, FES participants produced responses that were as semantically interrelated as controls' responses (AI). Both LTLE and FES participants accessed fewer clusters compared to controls, but once a semantic cluster was accessed, FES participants generated as many related exemplars as controls (cluster size). By contrast, LTLE participants generated smaller clusters relative to controls. Analyses of the percent of perseverative responses showed FES participants generated a higher rate of perseverative errors relative to LTLE participants and controls. Finally, FES participants' ANWLG total responses significantly correlated with tests of both language and executive functions, while only the correlation between total responses and a language test (BNT) was significant in the LTLE group.

Based on our findings that (1) FES and LTLE participants generated significantly fewer ANWLG responses relative to controls, and (2) the FES group produced responses that were as semantically related (AI) as control participants, we suggest that FES participants possess relatively preserved semantic knowledge and that semantic knowledge deficits do not account for reduced ANWLG in FES.

Table 4. Animal fluency and neuropsychological test correlations

Measure	First episode schizophrenia			Left temporal lobe epilepsy			Controls		
	N	*r*	*p*	*N*	*r*	*p*	*N*	*r*	*p*
Language									
Sentence Repetition	*47*	*.44*	*.002*	50	.33	.019	30	−.09	.639
Token Test*	47	.41	.004	10	.58	.079	30	.38	.037
Boston Naming Test	*47*	*.62*	*<.001*	*56*	*.42*	*.001*	29	.17	.376
Executive functions									
Trails–A	*47*	*−.46*	*.001*	56	−.16	.250	31	−.38	.035
Trails–B	*47*	*−.45*	*.002*	55	−.17	.208	31	−.50	.005
WCST–Categories*	47	.39	.007	54	.20	.151	31	.21	.250
WCST–Perseverations*	47	−.35	.016	54	−.30	.032	31	−.21	.259

*Pearson Rank Order correlations. *Note.* Bonferroni corrected p value = .05/21 = .002.
Significant correlations are in italics.

Our results differ from past studies that have reported semantic knowledge degradation in schizophrenia based on ANWLG performance (Aloia et al., 1996; Moelter et al., 2001; Paulsen et al., 1996; Robert et al., 1998; Rossell et al., 1999). This discrepancy may be explained by differences in methodology, as several past studies have used procedures that analyzed only a subset of the most common responses, which included a larger percent of control responses than patient responses (Aloia et al., 1996; Paulsen et al., 1996; Rossell et al., 1999). The AI, used in the present study to assess the semantic organization of ANWLG responses, explores the semantic relatedness of all of the responses produced by each participant. Thus, such semantically rich response strings as *wolf, jackal, hyena,* which are not captured with multidimensional scaling techniques, are quantified with the AI.

Our study also differed from past studies of chronic patients, such that we included only schizophrenia patients who had been hospitalized for a single episode of psychotic illness and who had received less than 12 weeks of cumulative neuroleptic treatment. Past studies showing semantic deficit in patients with schizophrenia have studied only chronic samples. Paulsen et al. (1996), however, did report a difference in the semantic organization of patients with late *versus* early onset schizophrenia and suggested that the difference may reflect disease progression, the long-term effect of decreased socialization, or premorbid neurodevelopmental differences. Our results strongly urge that further study is necessary to explore the relationship between disease chronicity and semantic knowledge processing in schizophrenia.

Finally, our study was also unique in that we included a LTLE comparison group with known temporal lobe damage and expected semantic knowledge deficits. The LTLE comparison group generated fewer ANWLG Total Responses and performed significantly more poorly on all ANWLG indices developed to assess semantic knowledge (i.e., AI, cluster size). This was not true for participants with FES. Thus, the comparison group served to validate our measures and support our claim that semantic knowledge is relatively preserved in FES.

The results that point to preserved semantic knowledge in FES do not, however, explain why these patients generate fewer responses on ANWLG relative to controls. Thus, the following question remains: what accounts for the reduced output of FES patients on ANWLG? To answer this question we turn to the model of Gruenewald and Lockhead (1980). Within the context of that model, the results that we have just reviewed suggest that FES participants are not impaired in generating words that are semantically interrelated (i.e., AI). Furthermore, unlike LTLE patients, once FES participants accessed a semantic subcategory, they generated as many words as controls from this subcategory (cluster size). Together, these results suggest that the semantic component of ANWLG is relatively unimpaired in FES. By contrast, FES patients differed from controls (and performed comparably to LTLE participants) in their ability to switch to a new semantic subcategory (cluster). That is, unlike controls who were able to shift from one subcategory to another, FES and LTLE participants sampled significantly fewer subcategories. Thus, on Gruenewald and Lockhead's (1980) model, we conclude that ANWLG output in FES is reduced due to the inability to switch from one semantic subcategory to another in their search for animal names (i.e., executive component).

The results of the present study parallel those obtained in a previous study of ANWLG in patients with "cortical" (i.e., Alzheimer's disease) and "subcortical" dementia due to ischemic vascular disease (Giovannetti et al., 1997). Cortical dementia participants performed similarly to the LTLE group in the present study (i.e., impaired AI, fewer clusters and reduced cluster size relative to aged-matched controls), while subcortical dementia participants performed more similarly to the FES group (i.e., preserved AI and fewer clusters relative to controls, but larger clusters relative to the cortical group). Paulsen et al. (1995) have shown that 50% of a sample of chronic schizophrenia participants ($N = 175$) obtained memory profiles consistent with the retrieval deficit observed in subcortical dementia, and, therefore, they have suggested the possibility that memory dysfunction in schizophrenia is a consequence of subcortical pathology. The present findings extend the results of Paulsen et al. (1995) by demonstrating that schizophrenia participants also obtain a "subcortical" profile on ANWLG. However, like Paulsen et al. (1995), we acknowledge that this "subcortical" pattern of performance is also consistent with cortical (i.e., prefrontal or diffuse) dysfunction.

Further insight into the deficit(s) underpinning reduced ANWLG in FES may be gleaned from the perseverative response rate measure and correlational analyses. The rate of perseverative responses was significantly higher in the FES group relative to the LTLE and control groups suggesting poor response monitoring and inefficient search strategies in FES. An explanation of perseveration and poor response monitoring is also consistent with the finding that FES participants sampled fewer semantic subcategories (i.e., clusters). That is, FES participants may have been perseverating, or stuck on a single or few semantic subcategories preventing them from generating more responses from other subcategories. Support for this account is found in the many studies that report increased perseveration and poor response monitoring in schizophrenia participants on a wide range of language and nonlanguage tasks (Barr et al., 1989; Bilder & Goldberg, 1987; Crider, 1997; Perry & Braff, 1998).

The correlation analyses, however, showed ANWLG total responses significantly correlated with both tests of language and executive function. Thus, a generalized, global neuropsychological deficit, rather than a specific executive deficit, must also be considered as a potential explanation for reduced ANWLG output in FES. Several investigators have recently reported a generalized neuropsychological deficit in schizophrenia (Bilder et al., 2000; Mohamed et al., 1999), and have suggested that this general deficit may explain poor performance on specific tests of executive functions (Laws, 1999).

CONCLUSIONS

In summary, the present results differ from category fluency studies that report degraded semantic processing in chronic schizophrenia patients. Our methods differed from past studies in that all ANWLG responses were analyzed and only first-episode schizophrenia patients were recruited. The interpretation of our results was facilitated by the inclusion of neurologically impaired comparison group with known temporal lobe damage and semantic knowledge deficits. We urge others to include such comparison groups in addition to healthy controls in future studies. Additional research is needed to determine whether the discrepancy between our results and past studies is due to differences in data analysis (i.e., AI *vs*. multidimensional scaling) or participant samples (i.e., FES *vs*. chronic patients), as these factors cannot be disambiguated in the present study. We also acknowledge the limitations of the AI, which, like multidimensional scaling, takes time to score and produces small between-group differences that may not be useful for clinical diagnosis. Furthermore, the AI is limited in quantifying semantic associations among responses that are not related to the attributes included in the scoring procedure (i.e., the association between *lions*, *tigers*, and *bears* to a well-known movie). These caveats notwithstanding, the present results suggest that semantic knowledge processing on ANWLG is relatively preserved in FES. Moreover, reduced ANWLG in FES may be best explained as impaired response monitoring, deficient search processes, and/or global cognitive dysfunction.

ACKNOWLEDGMENTS

This research was initiated while all of the authors were affiliated with Long Island Jewish Medical Center and Hillside Hospital, New Hyde Park, NY. A portion of this research was presented at the annual meeting of the International Neuropsychological Society, Chicago, IL, 2001. We thank Dr. Igor Grant and three anonymous reviewers for their helpful feedback.

REFERENCES

Allen, H.A., Liddle, P.F., & Frith, C.D. (1993). Negative features, retrieval processes and verbal fluency in schizophrenia. *British Journal of Psychiatry*, *163*, 769–775.

Aloia, M.S., Gourovitch, M.L., Weinberger, D.R., & Goldberg, T.E. (1996). An investigation of semantic space in patients with schizophrenia. *Journal of the International Neuropsychological Society*, *2*, 267–273.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Barr, W.B., Bilder, R.M., Goldberg, E., Kaplan, E., & Mukhergee, S. (1989). The neuropsychology of schizophrenic speech. *Journal of Communication Disorders*, *22*, 327–349.

Bell, B.D., Hermann, B.P., Woodard, A.R., Jones, J.E., Rutecki, P.A., Sheth, R., Dow, C.C., & Seidenberg, M. (2001). Object naming and semantic knowledge in temporal lobe epilepsy. *Neuropsychology*, *15*, 434–443.

Benton, A.L. & Hamsher, K. deS. (1978). *Multilingual Aphasia Examination*. Iowa City: University of Iowa.

Bilder, R.M., Bogerts, B., Ashtari, M., Wu, H., Alvir, J.M., Jody, D., Reiter, G., Bell, L., & Lieberman, J. A. (1995). Anterior hippocampal volume reductions predict "frontal lobe" dysfunction in first episode schizophrenia. *Schizophrenia Research*, *17*, 47–58.

Bilder, R.M. & Goldberg, E. (1987). Motor perseverations in schizophrenia. *Archives of Clinical Neuropsychology*, *2*, 195–214.

Bilder, R.M., Goldman, R.S., Robinson, D., Reiter, G., Bell, L., Bates, J.A., Pappadopulos, E., Willson, D.F., Alvir, J., Woerner, M.G., Geisler, S., Kane, J.M., & Lieberman, J.A. (2000). Neuropsychology of first episode schizophrenia: Initial characterization and clinical correlates. *American Journal of Psychiatry*, *157*, 549–559.

Bilder, R.M., Lipschutz-Broch, L., Reiter, G., Geisler, S., Mayerhoff, D., & Lieberman, J.A. (1991). Neuropsychological deficits in the early course of first episode schizophrenia. *Schizophrenia Research*, *5*, 198–199.

Bilder, R.M., Lipschutz-Broch, L., Reiter, G., Geisler, S.H., Mayerhoff, D.I., & Lieberman, J.A. (1992). Intellectual deficits in first-episode schizophrenia: Evidence for progressive deterioration. *Schizophrenia Bulletin*, *18*, 437–448.

Chan, A.S., Butters, N., Paulsen, J.S., Salmon, D.P., Swenson, M.R., & Maloney, L.T. (1993). An assessment of the semantic network in patients with Alzheimer's disease. *Journal of Cognitive Neuroscience*, *5*, 254–261.

Chertkow, H., Bub, D., Deaudon, C., & Whitehead, V. (1997). On the status of object concepts in aphasia. *Brain and Language*, *58*, 203–232.

Crider, A. (1997). Perseveration in schizophrenia. *Schizophrenia Bulletin*, *23*, 63–74.

DeRenzi, E. & Falgoni, P. (1978). Normative data and screening power of a shortened version of the Token Test. *Cortex*, *14*, 41–49.

Friedman, J.I., Harvey, P.D., Kemether, E., Byne, W., & Davis, K. (1999). Cognitive and functional changes with aging in schizophrenia. *Biological Psychiatry*, *46*, 921–928.

Frith, C.D. (1992). *The cognitive neuropsychology of schizophrenia*. Hove, UK: Erlbaum.

Garrard, P., Perry, R., & Hodges, J. (1997). Disorders of semantic memory. *Journal of Neurology, Neurosurgery and Psychiatry*, *62*, 431–435.

Giovannetti-Carew, T., Lamar, M., Cloud, B.S., Grossman, M., & Libon, D.J. (1997). Impairment in category fluency in ischemic vascular dementia. *Neuropsychology*, *11*, 400–412.

Gold, J.M., Blaxton, T.A., Hermann, B.P., Randolph, C., Fedio, P., Goldberg, T.E., Theodore, W.H., & Weinberger, D.R. (1994). Memory and intelligence in lateralized temporal lobe epilepsy and schizophrenia. *Schizophrenia Research*, *17*, 59–65.

Goldberg, T.E., Aloia, M.S., Gourovitch, M.L., Missar, D., Pickar, D., & Weinberger, D.R. (1998). Cognitive substrates of thought disorder, I: The semantic system. *American Journal of Psychiatry*, *155*, 1671–1676.

Goldstein, R.Z., Giovannetti, T., Schullery, M., Zuffante, P.A., Lieberman, J.A., Robinson, D.G., Barr, W.B. & Bilder, R.M. (2002). Predicting formal thought disorder from semantic vs. executive dysfunction in first-episode schizophrenia. *Journal of Neuropsychology, Neuropsychiatry and Behavioral Neurology*, *15*, 88–89.

Gourovitch, M.L., Goldberg, T.E., & Weinberger, D.R. (1996). Verbal fluency deficits in patients with schizophrenia: Seman-

tic fluency is differentially impaired as compared with phonologic fluency. *Neuropsychology*, *10*, 573–577.
Grant, D.A. & Berg, E.A. (1948). A behavioral analysis of the degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem. *Journal of Experimental Psychology*, *38*, 404–411.
Gruenewald, P.J. & Lockhead, G.R. (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 225–240.
Hart, J. & Gordon, B. (1990). Delineation of single-word semantic comprehension deficits in aphasia, with anatomical correlation. *Annals of Neurology*, *27*, 226–231.
Helmstaedter, C., Gleissner, U., Di Perna, M., & Elger, C.E. (1997). Relational verbal memory processing in patients with temporal lobe epilepsy. *Cortex*, *33*, 667–678.
Joyce, E.M., Collinson, S.L., & Crichton, P. (1996). Verbal fluency in schizophrenia: Relationship with executive function, semantic memory and clinical alogia. *Psychological Medicine*, *26*, 39–49.
Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test* (2nd ed.). Philadelphia: Lea & Febiger.
Laws, K.R. (1999). A meta-analytic review of Wisconsin Card Sort studies in schizophrenia: General intellectual deficit in disguise? *Cognitive Neuropsychiatry*, *4*, 1–30.
Lieberman, J.A., Alvir, J.M., Woerner, M., Degreef, G., Bilder, R.M., Ashtari, M., Bogerts, B., Mayerhoff, D.I., Geisler, S.H., Loebel, A., Levy, D.L., Hinrichsen, G., Szymanski, S., Chakos, M., Koreen, A., Borenstein, M., & Kane, J.M. (1992). Prospective study of psychobiology in first-episode schizophrenia at Hillside hospital. *Schizophrenia Bulletin*, *18*, 351–371.
Lieberman, J.A., Jody, D., Alvir, J.M., Ashtari, M., Levy, D.L., Bogerts, B., Degreef, G., Mayerhoff, D.I., & Cooper, T. (1993b). Brain morphology, dopamine, and eye-tracking abnormalities in first-episode schizophrenia. *Archives of General Psychiatry*, *50*, 357–368.
Lieberman, J.A., Jody, D.N., Geisler, S.H., Alvir, J.M., Loebel, A., Szymanski, S., Woerner, M., & Borenstein, M. (1993a). Time course and biologic correlates of treatment response in first-episode schizophrenia. *Archives of General Psychiatry*, *50*, 369–376.
McKay, A.P., McKenna, P.J., Bentham, P., Mortimer, A.M., Holbery, A., & Hodges J.R. (1996). Semantic memory is impaired in schizophrenia. *Biological Psychiatry*, *39*, 929–937.
Moelter, S.T., Hill, S.K., Ragland, J.D., Lunardelli, A., Gur, R.C., Gur, R.E., & Moberg, P.J. (2001). Executive and semantic system impairment during animal word list generation in schizophrenia. *Neuropsychology*, *15*, 502–509.
Mohamed, S., Paulsen, J.S., O'Leary, D., Arndt, S., & Andreasen, N. (1999). Generalized cognitive deficits in schizophrenia: A study of first-episode patients. *Archives of General Psychiatry*, *56*, 749–754.
Papps, B.P., Best, J.J.K., & O'Carroll, R.E. (2000). Semantic memory functioning and the left temporal lobe. *Neurocase*, *6*, 179–192.
Paulsen, J.S., Romero, R., Chan, A., Davis, A.V., Heaton, R.K., & Jeste, D.V. (1996). Impairment of the semantic network in schizophrenia. *Psychiatry Research*, *63*, 109–121.
Paulsen, J.S., Heaton, R.K., Sadek, J.R., Perry, W., Delis, D.C., Braff, D., Kuck, J., Zisook, S., & Jeste, D.V. (1995). The nature of learning and memory impairments in schizophrenia. *Journal of the International Neuropsychological Society*, *1*, 88–99.
Perry, W. & Braff, D.L. (1998). A multimethod approach to assessing perseverations in schizophrenia patients. *Schizophrenia Research*, *33*, 69–78.
Rossell, S.L., Rabe-Hesketh, S., Shapleske, J., & David, A.S. (1999). Is semantic fluency differentially impaired in schizophrenia patients with delusions? *Journal of Clinical and Experimental Neuropsychology*, *21*, 629–642.
Robert, P.H., Lafont, V., Medecin, I., Berthet, L., Thauby, S., Baudu, C., & Darcourt, G. (1998). Clustering and switching strategies in verbal fluency tasks: Comparison between schizophrenics and healthy adults. *Journal of the International Neurological Society*, *4*, 539–546.
Robinson, D., Woerner, M.G., Alvir, J.M., Bilder, R., Goldman, R., Geisler, S., Koreen, A., Sheitman, B., Chakos, M., Mayerhoff, D., & Lieberman, J.A. (1999a). Predictors of relapse following response from a first episode of schizophrenia or schizoaffective disorder. *Archives of General Psychiatry*, *56*, 241–247.
Robinson, D., Woerner, M.G., Alvir, J.M., Geisler, S., Koreen, A., Sheitman, B., Chakos M., Mayerhoff, D., Bilder, R., Goldman, R., & Lieberman, J.A. (1999b). Predictors of treatment response from a first episode of schizophrenia or schizoaffective disorder. *American Journal of Psychiatry*, *156*, 544–549.
Seidman, L.J., Stone, W.S., Jones, R., Harrison, R.H., & Mirsky, A.F. (1998). Comparative effects of schizophrenia and temporal lobe epilepsy on memory. *Journal of the International Neuropsychological Society*, *4*, 342–352.
Spitzer, R.L., Endicott, J., & Robins, E. (1978). Research Diagnostic Criteria: Rationale and reliability. *Archives of General Psychiatry*, *35*, 773–782.
Spreen, O. & Benton, A.L. (1969). *Neurosensory Center Comprehensive Examination for Aphasia*. Victoria, British Columbia, Canada: Neuropsychological Laboratory Department of Psychology, University of Victoria.
Tröster, A.I., Warmflash, V., Osorio, I., Paolo, A.M., Alexander, L.J., & Barr, W.B. (1995). The roles of semantic networks and search efficiency in verbal fluency performance in intractable temporal lobe epilepsy. *Epilepsy Research*, *21*, 19–26.
Troyer, A.K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology*, *22*, 370–378.
Troyer, A.K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology*, *11*, 138–146.
Troyer, A.K., Moscovitch, M., Winocur, G., Alexander, M.P., & Stuss, D. (1998a). Clustering and switching on verbal fluency: The effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia*, *36*, 499–504.
Troyer, A.K., Moscovitch, M., Winocur, G., Leach, L., & Freedman, M. (1998b). Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease. *Journal of the International Neuropsychological Society*, *4*, 137–143.
Warrington, E.K. & Shallice, T. (1979). Semantic access dyslexia. *Brain*, *102*, 43–63.
Wechsler, D. (1987). Wechsler Adult Intelligence Scale–Revised. San Antonio, TX: The Psychology Corporation.
Zakzanis, K.K., Troyer, A.K., Rich, J.B., & Heinrichs, W. (2000). Component analysis of verbal fluency in patients with schizophrenia. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *13*, 239–245.