JSLHR

Research Article

Weaknesses in Lexical-Semantic Knowledge Among College Students With Specific Learning Disabilities: Evidence From a Semantic Fluency Task

Jessica Hall,[a] Karla K. McGregor,[a] and Jacob Oleson[a]

Purpose: The purpose of this study is to determine whether deficits in executive function and lexical-semantic memory compromise the linguistic performance of young adults with specific learning disabilities (LD) enrolled in postsecondary studies.
Method: One hundred eighty-five students with LD (n = 53) or normal language development (ND, n = 132) named items in the categories *animals* and *food* for 1 minute for each category and completed tests of lexical-semantic knowledge and executive control of memory. Groups were compared on total names, mean cluster size, frequency of embedded clusters, frequency of cluster switches, and change in fluency over time. Secondary analyses of variability within the LD group were also conducted.
Results: The LD group was less fluent than the ND group. Within the LD group, lexical-semantic knowledge predicted semantic fluency and cluster size; executive control of memory predicted semantic fluency and cluster switches. The LD group produced smaller clusters and fewer embedded clusters than the ND group. Groups did not differ in switching or change over time.
Conclusions: Deficits in the lexical-semantic system associated with LD may persist into young adulthood, even among those who have managed their disability well enough to attend college. Lexical-semantic deficits are associated with compromised semantic fluency, and the two problems are more likely among students with more severe disabilities.

In the United States, students with specific learning disabilities (LD) are enrolling in postsecondary studies in record numbers (Sanford et al., 2011). Despite performing well enough academically to gain admission to a college or university, these students report greater difficulty with assignments and more obstacles imposed by their skill levels than other students (McGregor et al., 2016). LD involves deficits in processes that underlie the comprehension and expression of spoken or written language (Individuals with Disabilities Education Act [IDEA], 2004). Therefore, to understand the obstacles that might impede the academic success of college students with LD, and ultimately to reduce those obstacles, we must first document the processes that are deficient. In this article, we consider two verbal memory processes: executive function and lexical-semantic knowledge.

Executive function refers to a number of processes that control cognition and action. Executive control of memory involves attention to the task, inhibition of irrelevant information, strategic planning and search, flexibility or switching behavior, and concurrent processing or working memory (Henry, Messer, & Nash, 2012). Children with developmental language impairments present with deficits in executive function (Kapa & Plante, 2015), including attention (Im-Bolter, Johnson, & Pascual-Leone, 2006; Marton, 2008), inhibition (Bishop & Norbury, 2005: Henry et al., 2012; Marton, 2008), planning (Henry et al., 2012; Marton, 2008), and working memory (Ellis Weismer, Evans, & Hesketh, 1999; Henry et al., 2012; Montgomery, Magimairaj & Finney, 2010). Some groups also find deficits in switching (Marton, 2008), but others do not (Im-Bolter et al., 2006). Executive function deficits also characterize those whose LD is manifest as a reading impairment rather than a language impairment (Johnson, Humphrey, Mellard, Woods, & Swanson, 2010).

Another area of verbal memory deficit involves the long-term lexical-semantic store. Specifically, problems with learning—or committing words to the long-term

[a]The University of Iowa, Iowa City
Correspondence to Jessica Hall: jessica-e-hall@uiowa.edu

Editor: Sean Redmond
Associate Editor: Janna Oetting
Received December 22, 2015
Revision received June 3, 2016
Accepted August 22, 2016
https://doi.org/10.1044/2016_JSLHR-L-15-0440

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

 Copyright © 2017 American Speech-Language-Hearing Association

store (Kan & Windsor, 2010)—as well as comprehending (Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998), naming (Kail & Leonard, 1986; McGregor, Newman, Reilly, & Capone, 2002), and defining (McGregor, Oleson, Bahnsen, & Duff, 2013) words already stored in long-term memory are symptomatic of developmental language impairments during childhood and adolescence. Again, deficits in long-term lexical-semantic knowledge may also characterize individuals with reading impairments, especially those impairments that involve poor reading comprehension (Elwér, Keenan, Olson, Byrne, & Samuelsson, 2013).

In the current study, we sought to determine whether deficits in executive function and long-term lexical-semantic memory continue to compromise linguistic performance in young adults with LD who have compensated for earlier symptomology well enough to pursue postsecondary studies. Becker and McGregor (2016) measured the amount of material that students with and without LD recalled from a college lecture. Those with LD recalled less. However, there was extensive variability within the LD group, and those who recalled less tended also to score more poorly on independent measures of verbal long-term and short-term memory. These findings suggest that residual deficits in verbal memory may pose obstacles to the academic success of some, but not all, college students with LD. The results motivate the current group-level analysis of verbal memory in college students as well as an analysis of variability within the group in terms of processes proposed to support performance (long-term lexical-semantic memory and executive function) as well as specific diagnoses and receipt of accommodations. The overall goal was to better understand underlying strengths and weaknesses influencing verbal memory performance in college students with LD.

We elected to use a semantic fluency task to reach this goal. Semantic fluency distinguishes school children with developmental language impairments from their unaffected age-mates (Henry et al., 2012; Weckerly, Wulfeck, & Reilly, 2001). Results for those with developmental reading impairments are mixed. Semantic fluency did not distinguish school children with reading impairment from their unaffected age-mates in Bental and Tirosh (2007), but it did distinguish young adults who were described as having "severe" dyslexia from young adults with normal reading abilities (Kinsbourne, Rufo, Gamzu, Palmer, & Berliner, 1991).

Semantic Fluency

The semantic fluency task is ideal for our purposes because it taps executive function (Fournier-Vicente, Larigauderie, & Gaonac'h, 2008; Hedden, Lautenschlager, & Park, 2005; Rosen & Engle, 1997; Unsworth, Spillers, & Brewer, 2010) and lexical-semantic knowledge (e.g., Hedden et al., 2005; Hughes & Bryan, 2002; Ruff, Light, Parker, & Levin, 1997; Unsworth et al., 2010) but minimizes confounds because it requires little metalinguistic awareness and no reading, writing, or demonstration of syntactic skill. Semantic fluency is not fully mature until the adult years (Welsh, Pennington, & Groisser, 1991); hence, the semantic fluency task should be sensitive to differences among young college students. Moreover, the semantic fluency task provides an efficient means of tapping exactly the sort of performance that is frequent in college classroom discussions and test situations—the generation of verbal information relevant to a given topic.

In the current study, we administered the semantic fluency task to college students with and without LD. Participants named all items from the categories *food* and *animals* that they could within 60 seconds. In accordance with the typical procedure, we measured fluency by counting the total number of names produced within the minute, minus any errors or repetitions. Below, we explain further opportunities for exploring verbal memory via the semantic fluency task. We specifically followed a long tradition of taking cluster behavior as an index of long-term lexical-semantic knowledge and switch behavior as an index of the executive control guiding strategic search (see summary and supporting experimental evidence in Troyer, Moscovitch, & Winocur, 1997).

Semantic Fluency as a Window Onto Lexical-Semantic Memory

Cluster Size

While naming across the minute-long time span of the semantic fluency task, people tend to produce semantically related clusters (e.g., *orange, grapefruit, tangerine*). Semantic clustering suggests knowledge of relationships between items (e.g., that oranges and grapefruits are types of fruit), as well as the organization of the semantic lexicon (e.g., that activation of orange spreads to the near neighbor grapefruit). In the words of Unsworth et al. (2010), "clustering reflects the propensity to traverse through the lexical-semantic store via associative linkages" (p. 452). Among typical adults, expressive vocabulary scores (and working memory performance) account for individual differences in mean cluster size; those with better vocabulary skills produce larger clusters (Unsworth et al., 2010). Older children, who presumably have larger and better organized lexicons, produce larger clusters than younger children (Sauzéon, Lestage, Raboutet, N'Kaoua, & Claverie, 2004).

Cluster Embedding

The relationships most likely tapped by a semantic fluency task are taxonomic coordinates and slot fillers. Taxonomic coordinates such as *ice cream* and *yogurt* are neighbors in a given neighborhood or category; in this case, the neighborhood is *dairy.* Slot fillers, such as *ice cream* and *cake* are members of a given script or event structure; here, the script is a *birthday party.* Taxonomic coordinates—and, to a lesser extent, slot fillers—exist at various levels of the hierarchy. In the case of the superordinate category of *food*, dairy is a primary cluster, and within dairy, ice creams are embedded. Thus, a listing such as *milk, cheese, chocolate ice cream, vanilla ice cream, strawberry ice cream, yogurt* is a dairy cluster with an ice cream cluster embedded within it. Similarly, a slot-filler listing such as *candy, ice cream, cake, pie,* *cookies* is a sweets cluster with a birthday cluster embedded within it.

The structure of primary clusters and embeddings can be determined by algorithms that calculate the frequency of the co-occurrence of any two items named and the distance between them within the entire list of names (Crowe & Prescott, 2003). However, this analysis is limited to a group-level description, and only the names listed by at least 15% of all participants can be included. In a master's thesis from our lab, Chen (2012) developed a coding procedure that enabled the use of the entire data set and determination of embeddings per individual participant. By applying this procedure, she found that 7-year-olds embedded more than 5-year-olds, who embedded more than 3-year-olds during semantic fluency tasks. This age-related difference in embedding suggests the development of hierarchical organization of the semantic lexicon, and relatedly, a depth of lexical-semantic knowledge that allows naming at subordinate (e.g., schnoodle) and superordinate (e.g., mammal) levels, as well as the earlier acquired basic level (e.g., dog; Mervis & Crisafi, 1982). We examined cluster size and cluster embedding to identify potential weaknesses in lexical-semantic depth and organization among the participants with LD.

Semantic Fluency as a Window Onto Executive Control of Memory

Cluster Switches

Naturally, not all of the words listed in a given semantic fluency task fit into a single cluster. Instead, people tend to switch from one cluster to another. Switches are of two sorts: hard switches, which are transitions between a cluster and a nonclustered word or between two nonclustered words; and cluster switches, which are transitions between clusters. In the listing *beef, bread, apple, peach, milk, cheese, ice cream,* there is a hard switch between *beef* and *bread* and another between *bread* and *apple.* The cluster switch between *peach* and *milk* marks a transition from fruit to dairy.

These distinctions are important because switch types reveal different processes. Abwender, Swan, Bowerman, and Connolly (2001) administered a semantic fluency task to healthy young adults and coded their switch behavior. They found that cluster switches were the better predictor of overall semantic fluency; that hard switches were associated with smaller semantic clusters; and that cluster switches, but not hard switches, predicted performance on an independent nonverbal measure of fluency. They interpreted this to mean that hard switches reveal difficulty in accessing semantic knowledge in long-term memory, whereas cluster switches reveal strategic search and flexibility. Raboutet et al. (2010) presented converging evidence but a slightly different interpretation; they also administered a semantic fluency task to healthy young adults, but they examined performance in the first and second halves of the minute-long interval. They found that hard switches were less frequent in the second half than the first but that cluster switches remained constant over time. Because the task becomes more effortful over time, they interpreted this to mean that hard switches reveal automatic retrieval processes (not difficulty in access), whereas cluster switches reveal more strategic searches of the long-term semantic store. Despite disagreements on the best interpretation of hard switching, both groups agree that hard switches and cluster switches are dissociable phenomena and that cluster switches are a sign of the strategic control of searching in the moment.

Change Over Time

The time segment approach of Raboutet et al. (2010) was motivated by the observation that people tend to retrieve many items in an initial burst and then fewer and fewer items over the course of the minute-long semantic fluency task. This change has been conceptualized as two stages: The first 15 seconds or so involves automatic activation of readily accessible words in the long-term store, whereas the final portion of the minute requires a more extensive, more controlled search involving executive aspects of short-term processing, such as strategy and inhibition (Crowe, 1998). Developmental support is available from a cross-sectional study of the semantic fluency of children ages 6 to 15 years (Hurks et al., 2010): The number of names generated in the first 15 seconds of the minute reached adult-like levels by age 10, while the number generated in the final 45 seconds of the minute reached adult-like levels two or more years later.

Because the demand for more controlled switching and monitoring increases as the minute unfolds, shallower declines over the minute-long interval are taken as a sign of stronger executive control over memory than steeper declines (Luo, Luk, & Bialystok, 2010). To illustrate, in a time course analysis of letter fluency responses, Luo et al. (2010) found that bilingual speakers demonstrated more gradual slopes of decline than monolingual speakers, which is consistent with the often reported bilingual advantage in executive functioning (see review in Bialystok, 2007). Following these examples, we examined two indices of change, cluster switching and number of items named within each time segment, to identify potential weaknesses in executive function among the participants with LD.

The Current Study

The primary purpose of this study was to explore patterns of semantic fluency that implicate executive control of memory and lexical-semantic memory as deficient processes among the students with LD. As preliminary steps, we also aimed to verify that college students with LD are less fluent than students with normal language development (ND) and to determine the extent to which individual differences in semantic fluency are associated with measures of executive function and lexical-semantic memory in these two groups of college students. We also determined whether semantic fluency within the LD group varied as a function of presence of ADHD, receipt of accommodations, and specific diagnosis.

We predicted the group with LD would exhibit lower semantic fluency than their peers with ND. Given previous literature, we predicted that cluster indices (cluster size and embedding) would correlate with scores on a measure of lexical-semantic knowledge, that change indices (cluster switches and change in naming over the minute) would correlate with scores on a measure of executive control of memory, and that overall semantic fluency would correlate with scores on both measures. We would interpret fewer words per cluster and less frequent embedding of clusters by adults with LD as evidence of weakness in long-term lexical-semantic memory and interpret fewer cluster switches and steeper slopes of change in number of words produced over the minute-long interval as evidence of weaknesses in executive control.

Method

Participants

Participants were adults, ages 18–25 years, who were currently enrolled in postsecondary education in the United States. The LD group comprised 53 students (29 women, 24 men), and the ND group comprised 132 students (68 women, 64 men). In the LD group, four students identified as Hispanic/Latino, 37 as not Hispanic/Latino, and 12 did not report their ethnic identity. One student identified as Asian, two as Black, 43 as White, six as more than one race, and one did not report racial identity. In the ND group, five students identified as Hispanic/Latino, 117 as not Hispanic/Latino, and 10 did not report ethnic identity. Six students identified as Asian, six as Black, 114 as White, five as more than one race, and one did not report racial identity. Groups were similarly composed in terms of type of postsecondary institution (community college, liberal arts college, or university), and they were well matched (*p*s > .50; Mervis & Robinson, 1999) on years of education, $t = 0.37$, $df = 182$, $p = .71$ (see Table 1 for detailed participant information). Seventy-four of the students also participated in McGregor (2014); 24 participated in McGregor, Arbisi-Kelm, and Eden (2016).

Participants considered for the LD group answered an advertisement asking for volunteers who "struggle with spoken or written language." To be included in the study, participants in the LD group met a classification criterion associated with a procedure developed by Fidler, Plante, and Vance (2011) to maximize sensitivity and specificity of the identification of developmental language impairments in adults. The procedure has been used with success by three independent research groups (Aguilar & Plante, 2014; McGregor, Arbisi-Kelm, & Eden, 2016; Poll, Watkins, & Miller, 2014). It involves administering language comprehension and spelling tasks, then weighting the score on each task according to its relative contribution to group discrimination. A positive overall score identifies students with a history of language impairments (i.e., those who received speech-language services as children) with a sensitivity of 80% and a specificity of 87% (see Table 6 in Fidler et al., 2011). All students with LD earned a positive score on the procedure of Fidler et al. (2011); all students with ND earned a negative score. Additional enrollment criteria were: a passing performance on a pure-tone hearing screening presented at 0.5, 1.0, 2.0, and 4.0 kHz at 20 dB bilaterally (American Speech-Language-Hearing Association, 1990); no positive history of acquired neurological disorders; and English as the primary language. All participants began to learn

Table 1. Participant demographics and test means (*M*s) and standard deviations (*SD*s) by diagnostic group (students with specific learning disabilities [LD] and students with normal language development [ND]) with effect sizes (Cohen's *d*) for significant between-groups differences.

Demographic/Test	LD	ND	Cohen's *d*
Age			
M	20.7	21.4	
SD	1.5	2.0	
min	18	18	
max	24	25	
Years of education			
M	13.7	14.8	
SD	1.4	1.9	
min	12	13	
max	17	21	
Percent reporting ADHD	25	.03	
Percent reporting language impairment	17.0	NA	
Percent reporting reading impairment	43.4	NA	
Percent reporting reading and language impairments	30.2	NA	
Percent unsure of official diagnosis	9.4	NA	
PPVT-4			
M	102	113	–.93*
SD	12	11	
min	80	86	
max	126	137	
EVT			
M	104	116	–.95*
SD	14	11	
min	75	89	
max	141	136	
NWR			
M	91	94	–.70*
SD	4.4	3.4	
min	79	86	
max	99	100	
CVLT-II			
M	45	55	–.96*
SD	9.5	9.6	
min	27	35	
max	69	79	
KBIT			
M	104	110	–.48*
SD	13	11	
min	80	86	
max	130	132	

Note. Scores on the Peabody Pictures Vocabulary Test–Fourth Edition (PPVT-4); Expressive Vocabulary Test (EVT); and Kaufman Brief Intelligence Test, nonverbal subtest (KBIT), are standard scores with a normative mean of 100 and a standard deviation of 15. Scores on Nonword Repetition (NWR) are a percentage of phonemes correct out of 96. Scores on the California Verbal Learning Test–Second Edition (CVLT-II) are *T* scores with a normative mean of 50 and a standard deviation of 10. min = minimum score; max = maximum score.

**p* < .02.

English before age 3 years; the majority reported learning English from birth.

We asked the 53 students who qualified for the LD group via the Fidler et al. (2011) protocol to report their diagnoses. It is not surprising that a number of diagnoses were reported, given the heterogeneity of the population (Leonard, 2014) and the variety of diagnostic labels that are applied to individuals within this population (Reilly, Bishop, & Tomblin, 2014). Thirty-seven participants were receiving accommodations from Student Disability Services on their campuses. Participants in the ND group had no prior or current diagnoses relevant to LD and no accommodations from Student Disability Services. We did not exclude students with attention deficit/hyperactivity disorder (ADHD), as this condition is often comorbid with LD (Mueller & Tomblin, 2012; see Table 1 for information on diagnoses).

To put this heterogeneity into perspective, it is essential to realize that access to postsecondary institutions in the United States is guaranteed by the Americans with Disabilities Act of 1990 (ADA; 1990), which specifies only that disabilities involve either "physical or mental impairment" and that a disability "substantially limits one or more major life activities" (Section 3.2.A). Unlike the Individuals with Disabilities Education Act (IDEA; 2004), which mandates a free, appropriate education for primary and secondary students with any of 13 recognized conditions, no specific diagnostic conditions are recognized under the ADA. Therefore, the heterogeneity of the group is true to the various manifestations of LD and to the legal view of the condition. That said, we did conduct secondary analyses to examine within-group variability according to reported accommodation status, LD diagnosis, and comorbidity of ADHD.

To more fully describe the participants, we administered the Peabody Picture Vocabulary Test–Fourth Edition (PPVT-4; Dunn & Dunn, 2007), the California Verbal Learning Test–Second Edition (CVLT-II; Delis, Kramer, Kaplan, & Ober, 2000), the Expressive Vocabulary Test (EVT; Williams, 2007), the nonverbal subtest of the Kaufman Brief Intelligence Test (K-BIT; Kaufman & Kaufman, 2004), and a nonstandardized test of nonword repetition (NWR; Dollaghan & Campbell, 1998). These tests were administered to all participants in the LD group. The PPVT-4 was administered to all participants in the ND group, but the other tests were administered to a subset of 54 individuals from the ND group. Within the ND group, the participants who received all tests were younger (M = 20.8 years, SD = 1.6) than those who received the PPVT-4 only (M = 21.7 years, SD = 2.1), $t = -2.6$, $df = 130$, $p = .01$, but these subgroups did not differ on PPVT-4 standard scores, $p = .78$. The LD group performed significantly lower than the ND group on all standardized tests of vocabulary and memory, but on average, their scores were not clinically significant (see Table 1).

Procedures

Participants were instructed to name as many items as possible in the categories *animals* and *food* in 1 minute for each category, with the order counterbalanced across participants. Examiners recorded all responses and informed participants when 60 seconds had expired.

Analysis

The examiners' written records were double checked against audio files, and the number of correct items, errors, and repetitions were determined. The files were then divided into four 15-second segments, and the number of correct items per segment was determined. We lacked audio files for one participant with LD in both *animals* and *food,* for one participant with ND in *animals,* and for another participant with ND in *food*; therefore, these data sets were excluded from the time segment analysis.

Clusters were determined using a combination of cluster types from Ross and Murphy (1999) and Troyer (2010), as well as natural co-occurrences in the data. Semantic clusters were of two varieties: taxonomic (items shared common traits or properties; e.g., a canine cluster or a dairy cluster), or slot-fillers (items shared a common context or script; e.g., a farm cluster or a birthday party cluster). Items were also coded for belonging to embedded clusters (e.g., types of cereal within a larger cluster of breakfast items).

We coded two types of switches: hard switches and cluster switches. Hard switches marked transitions between a response in a cluster and a response not in a cluster, or vice versa. Hard switches also occur between each item in a series of unclustered items. In contrast, cluster switches are transitions between two adjacent clusters. Table 2 provides examples of cluster and switch coding. Two independent coders rated transcripts and achieved point-to-point agreement above 90% for all coding tasks in a randomly selected 10% of transcripts that they both coded.

After coding was completed, dependent variables from each transcript were tallied per semantic category as follows: the total number of within-category names, repetitions, and errors (e.g., *marzipan* listed as an animal); total within-category names per 15-s time segment; number of clusters; mean cluster size; ratio of embedded clusters to total clusters of three or more items; and ratio of cluster switches to hard switches. Mean cluster size was determined by dividing the sum of the number of names in each cluster by the total number of clusters. Errors and repetitions were not counted in cluster size, but they were counted in tallies of switch behavior. Embedded clusters and cluster switches were expressed as ratios to correct for differences in the number of clusters that each person named. Switch behavior was calculated for main clusters only; embedded clusters were not considered.

To determine whether the LD and ND groups differed on any fluency indices, we ran multiple mixed analyses of variance (ANOVAs) with diagnostic group (LD, ND) and gender (male, female) as between-subjects factors and category type (*animals, food*) as a within-subject factor. Gender was included as a factor in recognition of the finding that lexical-semantic knowledge differs as a function of the interaction between gender and category (Capitani, Laiacona, &

Table 2. Examples of coded responses for *animal* and *food* categories.

Category	Response	Main cluster	Type	Embedded cluster	Type
Animals	bird	predator–prey	slot-filler		
	worm				
	spider	insect	taxonomic		
	ladybug				
	frog	water	slot-filler	reptiles/amphibians	taxonomic
	toad				
	penguin				
Food	celery	vegetable	taxonomic		
	cucumber				
	cheese	none			
	watermelon	fruit	taxonomic		
	cantaloupe				
	honeydew				
	bread	breads/grains/cereals	taxonomic		
	bagel			breakfast	slot-filler
	waffle				

Barbarotto, 1999) and the conclusion that women and men bring different processing strategies to bear on fluency tasks (Weiss et al., 2006).

To evaluate change in naming over time, we used a linear regression model with correlated errors to account for repeated observations per participant. General linear models make the assumption that errors, or residuals of the model, will not be correlated, but time-based models violate this assumption because the errors in one segment are correlated with errors in a second segment. A correlated errors regression model with an unstructured covariance matrix allows us to control for the correlation of errors between time segments. The best fitting model was chosen according to the Bayesian information criterion and included the factors diagnostic group, category, and time segment, as well as the interaction between segment and category. Analyses were generated using SAS version 9.3.

To determine the association between lexical-semantic knowledge and semantic fluency performance, we correlated raw scores on the PPVT-4 with the fluency indices: overall fluency, mean cluster size, ratio of embedded clusters to clusters of three or more items, and ratio of cluster switches to hard switches. Because the PPVT-4 requires recognition, not recall, of words and their semantic referents, it is an ideal independent measure of the long-term lexical-semantic knowledge demands associated with the semantic fluency task. To determine the association between executive function and semantic fluency performance, we correlated raw scores on Trial 1 of the CVLT-II to these same indices. Trial 1 involves immediate verbal recall of a list of 20 words that the examiner has presented orally. The words are in random order, but four different semantic categories are represented: *furniture*, *vegetables*, *transport*, and *animals*. Whereas the CVLT-II provides a broad measure of verbal memory, Trial 1 is more specifically sensitive to executive control of short-term memory. Performance on Trial 1 correlates moderately or strongly with aspects of executive function as measured by the Wechsler Memory Scale–Revised (Wechsler, 1987), including the attention/concentration index, mental control, and digits forward and backward (Delis, Cullum, Butters, Cairns, & Prifitera, 1988). Trial 1 of the CVLT-II is also an ideal choice because the stimuli are real words from semantic categories, like the naming responses on the semantic fluency task. Although our intention was to use the PPVT-4 to tap long-term lexical-semantic knowledge and Trial 1 of the CVLT-II to tap executive control of memory, neither is likely a pure measure of the respective constructs. However, a lack of correlation between the two test scores for the LD group ($r = .18$, $p = .24$) and the ND group ($r = -.006$, $p = .97$) suggests that we successfully selected measures that are independent.

Results

Fluency

Incorrect responses were rare. On average, the LD group produced 0.53 repetitions ($SE = 0.10$) and 0.17 errors ($SE = 0.06$) per category, and the ND group produced 0.46 repetitions ($SE = 0.06$) and 0.20 errors ($SE = 0.03$) per category. Given that these response types were near floor, we did not conduct any statistical analyses to explore these variables.

The LD and ND groups differed as predicted in overall fluency. A mixed ANOVA revealed a main effect of diagnostic group, $F(1, 181) = 6.64$, $p = .01$, partial $\eta^2 = .04$, such that LD participants named fewer items (see Table 3). There was also a main effect of category, $F(1, 181) = 15.06$, $p = .0001$, partial $\eta^2 = .08$, with fewer animal ($M = 24.68$, $SE = 0.52$) than food names ($M = 26.54$, $SE = 0.54$). There was neither a main effect of gender, nor were there interactions between diagnostic group, category, and gender.

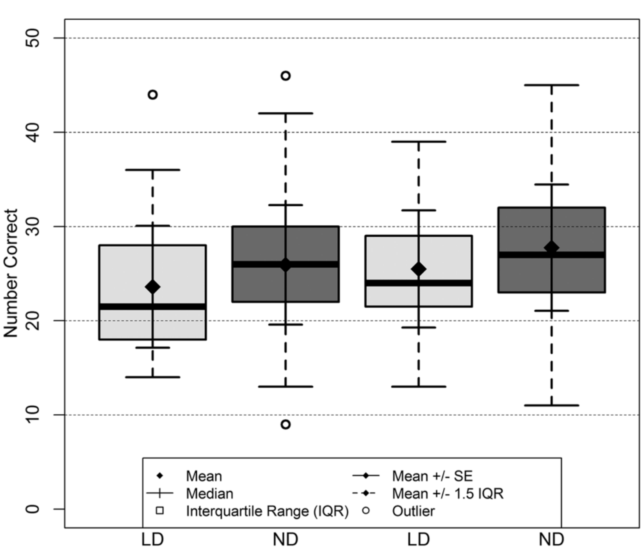
Despite the significant group-level difference, the overlap between the fluency of one group and the other was extensive, even at lower levels of performance (see Figure 1). This was especially true of the food category, in which 17% of the LD participants and 15% of the ND participants fell more than one standard deviation below

Table 3. Comparison of fluency indices by diagnostic group expressed as mean (*M*), standard error (*SE*), minimum score (min) and maximum score (max) with effect sizes for significant between-groups differences expressed as partial η^2.

Fluency index	LD	ND	Partial η^2
Overall fluency			
M	24	27	.04*
SE	0.80	0.50	
min	13	9	
max	44	46	
Cluster size			
M	3.18	3.42	.02*
SE	0.10	0.06	
min	2	2	
max	6.4	12.25	
Embedded cluster ratio			
M	.39	.54	.05*
SE	.03	.03	
min	0	0	
max	1	1.63	
Cluster switch ratio			
M	.40	.39	*ns*
SE	.02	.01	
min	0	0	
max	1	1	
Slope			
M	5.22	5.44	*ns*
SE	0.32	0.20	
min	–1	–4	
max	15	17	

*$p < .05$.

Figure 1. The mean number of correct names by category (*animals* and *foods*) and diagnostic group (students with specific learning disabilities [LD] and those with normal language development [ND]).



the mean performance of the ND group. For the *animal* category, 35% of the LD participants and 13.5% of the ND participants fell more than one standard deviation below the mean performance of the ND group.

Clustering Behavior

Semantic clusters were used by all participants. The LD group averaged 6.9 food clusters (SE = 0.29), and 77% of all foods named were clustered. For animals, they averaged 6.8 clusters (SE = 0.29), and 86% of all animals named were clustered. The ND group averaged 6.8 food clusters (SE = 0.20), and 78% of foods named were clustered. They averaged 7.2 animal clusters (SE = 0.18), and 85% of animals named were clustered. Despite these similarities, a mixed ANOVA with mean cluster size as the dependent variable revealed a main effect of diagnostic group, $F(1, 181) = 4.02$, $p = .046$, partial $\eta^2 = .02$. The LD group had smaller clusters than the ND group (see Table 3). There was no main effect of gender or category, and there were no interactions between diagnostic group, category, and gender.

With the ratio of embedded clusters to total clusters of three or more items as the dependent variable, a mixed ANOVA revealed a main effect of diagnostic group, $F(1, 177) = 10.14$, $p = .002$, partial $\eta^2 = .05$. The LD group demonstrated a lower rate of embedded clusters than the ND group (see Table 3). There were no main effects involving category or gender, and there were no interactions.

Change Behavior

With the ratio of cluster switches to hard switches as the dependent variable, a mixed ANOVA revealed no main effect of diagnostic group (see Table 3). There was a main effect of category, $F(1, 181) = 26.62$, $p < .0001$, partial $\eta^2 = .13$, with a higher ratio for the *animal* category (M = 0.46, SE = 0.02) than the *food* category (M = 0.33, SE = 0.02); that is, participants used relatively more cluster switches when naming animals than when naming foods. There were no main effects of gender, and there were no significant interactions.

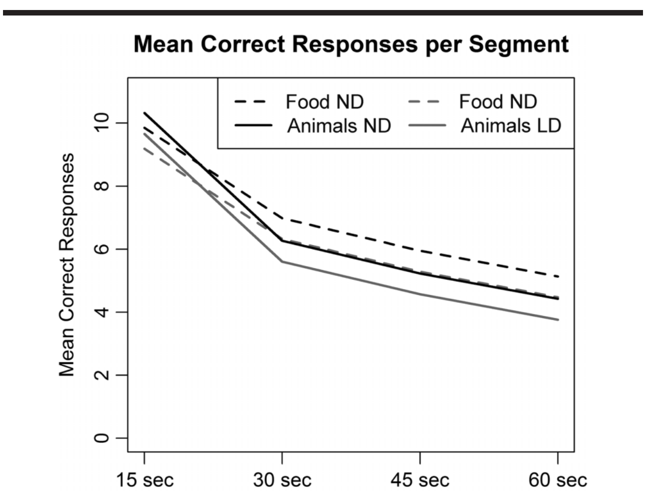
One may question whether cluster switches are always superior to hard switches. A particularly inefficient pattern of searching would be one in which given cluster types are briefly explored, abandoned, and then returned to repeatedly during the minute-long interval. This would inflate the ratio of cluster switches to hard switches. To ensure that the lack of difference between the LD and ND groups in the ratio of cluster switches to hard switches did not mask a greater dependence on repeated clusters on the part of the LD group, we ran a mixed ANOVA with number of repeated clusters as the dependent variable. There were no significant interactions. There was no main effect of diagnostic group (LD, M = 1.05, SE = 0.09; ND, M = 1.11, SE = 0.06). There was a main effect of category, $F(1, 181) = 14.96$, $p < .001$, partial $\eta^2 = .08$, with fewer repetitions within the *animal* category (M = 0.88, SE = 0.07) than in the *food* category (M = 1.28, SE = 0.08). There was also a main effect of gender, $F(1, 181) = 10.36$, $p = .002$, partial $\eta^2 = .05$, such that women (M = 1.26, SE = 0.07) repeated more often than men (M = 0.91, SE = 0.08). In summary, we found no evidence that the LD and ND groups differed in use of cluster switches relative to hard switches.

To analyze change over the minute-long interval, we first conducted a mixed ANOVA with slope (number of items in Time Segment 1 minus number of items in Time Segment 4) as the dependent variable. There was no main effect of diagnostic group, and there were no interactions involving diagnostic group. There was no main effect of gender, but there was a main effect of category, $F(1, 180) = 19.89$, $p = .00001$, partial $\eta^2 = .10$, that was qualified by a Category × Gender interaction, $F(1, 180) = 4.55$, $p = .03$, partial $\eta^2 = .02$. A Bonferroni post hoc test revealed that women demonstrated a steeper slope for animals (M = 6.6, SE = 0.32) than for foods (M = 4.5, SE = 0.34), $p < .0001$; among men, slopes for animals (M = 5.5, SE = 0.36) and foods (M = 4.7, SE = 0.38) did not differ, $p = .46$

Because performance in the first 15 seconds of the task is interpreted differently than performance in the final 45 seconds (Crowe, 1998; Hurks et al., 2010; Luo et al., 2010), a more nuanced analysis by time segment was required. To determine whether there were different patterns of change over time by diagnostic group or semantic category, we analyzed semantic fluency by 15-second time segments (see Figure 2). Residual plots showed no violation of model assumptions.

As we knew from the previous analysis, the difference between diagnostic groups was significant, with the LD group naming fewer items than the ND group, $t(182) = -2.99$, $p = .003$, as was the difference between categories, with more foods named than animals named, $t(182) = -3.53$, $p = .0005$. Critically, there was no interaction between group and category and, therefore, no evidence that the groups differed in change over time.

Figure 2. Model-predicted number of correct responses produced per 15-second segment by diagnostic group (students with specific learning disabilities [LD] and those with normal language development [ND]) and category (*animals* and *foods*).



Not surprisingly, we found a main effect of time segment such that the average rate of production regardless of group or category lowered across the 15-second time segments, $F(3, 182) = 391.96$, $p < .0001$, with more items named in the first segment than in the second, more in the second than in the third, and more in the third than in the fourth, in both categories, $p < .001$ for all comparisons. There was a main effect of category, $F(1, 182) = 15.55$, $p < .0001$, that was qualified by a significant interaction between segment and category, $F(3, 182) = 9.14$, $p < .0001$, meaning that the difference between number of animals and foods named depends on which time segment we focus on. Breaking these interactions apart, participants named fewer foods than animals in the first segment, $t(182) = 2.25$, $p = .03$; and more foods than animals in segments two, $t(182) = -3.74$, $p = .0002$; three, $t(182) = -3.53$, $p = .0005$; and four, $t(182) = -3.53$, $p = .0005$.

Individual Differences

Tables 4 and 5 present correlation matrices for the fluency indices and the measures of lexical-semantic knowledge and executive control of memory. First note that, for both the LD and ND groups, overall fluency was positively related to the cluster switch to hard switch ratio and mean cluster size, as would be expected, given that many clusters and large clusters should enhance overall productivity. For the ND group only, cluster size and the ratio of embedded clusters to total clusters of three or more items was related. This too is logical, given that larger clusters can contain more embeddings. This correlation was not significant for the LD group, but recall that their mean cluster size was smaller than that of the ND group; that is, they had a more restricted range of cluster sizes than the ND group.

We now turn to the predicted relationships. We predicted that (a) both lexical-semantic knowledge and executive control of memory would be positively associated with overall fluency; (b) lexical-semantic knowledge in particular would be positively associated with the cluster indices; and (c) executive control of memory would be positively associated with the change indices. These predictions held, in part, for the LD group (see Table 4). Scores on the PPVT-4, our measure of lexical-semantic knowledge, correlated with overall fluency and cluster size. There was no relationship between PPVT-4 and embedded cluster ratio, but again, there were a restricted number of possibilities for embedding in this group given their small cluster sizes. Scores on Trial 1 of the CVLT-II, our measure of executive control of memory, correlated with overall fluency and the cluster to hard switch ratio. There was no correlation between Trial 1 of the CVLT-II and slope of change over the course of the minute.

None of the predictions held for the ND group, with the exception of a small, positive association between the PPVT-4 and overall fluency (see Table 5). This was a more

Table 4. Correlations between measures of lexical-semantic knowledge (PPVT-4) and executive control of memory (CVLT-II, Trial 1) and fluency indices for the students with significant learning disabilities.

Test	PPVT-4	CVLT-II, Trial 1	Overall fluency	Cluster size	Embedded cluster ratio	Cluster switch ratio	Slope
PPVT-4	1.0						
CVLT-II, Trial 1	.18	1.0					
Overall fluency	.40**	.45**	1.0				
Cluster size	.35*	.26	.56**	1.0			
Embedded cluster ratio	−.04	.14	.07	.16	1.0		
Cluster switch ratio	.00	.35*	.31*	.12	.20	1.0	
Slope	−.25	−.01	.14	−.27	.03	.24	1.0

Note. PPVT-4 = Peabody Picture Vocabulary Test–Fourth Edition, CVLT-II = California Verbal Learning Test–Second Edition. $n = 50$.
*$p < .05$. **$p < .01$.

modest correlation than in the LD group, but given the larger sample size of the ND group, it was significant.

Given the single small correlation and the many null results in the ND group, we examined the distribution of scores in scatter plots. The scores earned by the ND group covered as large or larger ranges as the scores earned by the LD group (see Tables 1 and 3). Also, these ranges did not include numerous ceiling- or floor-level performances. The largest number of such performances within the ND group occurred when naming foods, for which seven participants used no cluster switches. Given that this represents only 5% of the ND group, floor-level performance was clearly not a frequent problem.

To further explore variability within the LD group, we analyzed performance according to the presence of ADHD, receipt of accommodations, and diagnosis. We ran a series of one-way ANOVAs to compare number of correct items named, cluster size, cluster embedding ratio, cluster switching ratio, and slope between subgroups within the LD group. In all analyses, we collapsed across gender and averaged across semantic categories. Finally, we ran a linear regression model with correlated errors to compare the subgroups' changes in naming over time segments. The crucial result would be an interaction between time segment and subgroup. However, for each subgroup analysis, a model with no interaction was the best fit; therefore, we will not present further details of the linear models.

First, we compared the 13 participants in the LD group who reported ADHD to the 40 who did not (see Table 6). We found no differences between these subgroups in overall fluency, cluster size, embedded cluster ratio, cluster switch ratio, or slope.

Next, we compared the 37 participants in the LD group who reported receipt of classroom accommodations to the 15 who did not (one participant did not provide information on accommodation; see Table 6). Those without accommodations named more items correctly than those with accommodations, $F(1, 50) = 5.02$, $p = .03$, partial $\eta^2 = .09$. These subgroups did not differ in cluster size, embedded clusters, cluster switching, or slope.

Finally, we compared the 23 participants with reading impairment, the nine participants with language impairment, and the 16 participants with both (see Table 6). Five other participants did not report or did not know their diagnoses, and they were excluded from this analysis. The ANOVA revealed a main effect of diagnostic group, $F(2, 45) = 4.07$, $p = .02$, partial $\eta^2 = .15$. Post hoc comparisons using Tukey's honestly significant difference (HSD) for unequal *n*s confirmed that participants with reading impairment only were more fluent than participants with both language and reading impairments, $p = .05$. The performance of participants with language impairment only fell between the other two subgroups and did not differ from either subgroup. We found the same pattern of results for mean cluster size,

Table 5. Correlations between measures of lexical-semantic knowledge (PPVT-4) and executive control of memory (CVLT-II, Trial 1) and fluency indices for the students with normal language development.

Test	PPVT-4	CVLT-II, Trial 1	Overall fluency	Cluster size	Embedded cluster ratio	Cluster switch ratio	Slope
PPVT-4	1.0						
CVLT-II, Trial 1	−.01	1.0					
Overall fluency	.18*	.17	1.0				
Cluster size	.15	−.05	.45**	1.0			
Embedded cluster ratio	.03	−.16	.32**	.47**	1.0		
Cluster switch ratio	.03	.19	.27**	.13	−.03	1.0	
Slope	−.14	.14	.08	−.08	−.04	−.02	1.0

Note. PPVT-4 = Peabody Picture Vocabulary Test–Fourth Edition, CVLT-II = California Verbal Learning Test–Second Edition. $n = 47$ for all correlations involving the CVLT, $n = 122$ for all other correlations.
*$p < .05$. **$p < .01$.

Table 6. Comparison of fluency indices by subdivisions within the students with specific learning disabilities (LD) group expressed as mean (*M*), standard error (*SE*), minimum score (min), and maximum score (max).

	Subdivisions of the LD group						
Fluency index	–ADHD (*n* = 40)	+ADHD (*n* = 13)	–Accom (*n* = 15)	+Accom (*n* = 37)	RI only (*n* = 23)	LI only (*n* = 9)	RI + LI (*n* = 16)
Overall fluency							
M	25	23	27[a]	23[a]	27[b]	23	22[b]
SE	0.89	1.6	1.7	0.77	1.1	1.58	1.2
min	15.5	16	18	15.5	15.5	19.5	16
max	36.5	35.5	36.5	36	36.5	32.5	35.5
Cluster size							
M	3.2	3.1	3.1	3.2	3.4[c]	3.0	2.9[c]
SE	0.09	0.12	0.13	0.09	0.10	.16	0.12
min	2.4	2.6	2.4	2.5	2.4	2.4	2.5
max	5.2	4.2	4.3	5.2	5.2	3.8	3.6
Embedded cluster ratio							
M	.39	.37	.43	.37	.44	.26	.38
SE	.04	.06	.06	.04	.05	.08	.06
min	0	.13	.10	0	.10	0	0
max	1.0	1.0	1.0	1.0	.89	.60	1.0
Cluster switch ratio							
M	.42	.33	.39	.40	.43	.36	.37
SE	.03	.05	.04	.03	.03	.05	.04
min	.04	.05	.04	.04	.20	.08	.04
max	.79	.69	.61	.79	.75	.65	.79
Slope							
M	5.1	5.7	5.1	5.2	5.5	5.1	5.4
SE	0.34	0.27	0.54	0.28	0.40	0.82	0.45
min	2.0	3.5	2.0	2.0	2.0	3.0	2.0
max	10.5	7.5	9.0	9.5	10.5	9.5	8.0

Note. ADHD = attention deficit/hyperactivity disorder; Accom = classroom accommodations; RI = reading impairment; LI = language impairment. Values marked by like superscripts differ at $p < .05$.

$F(2, 45) = 5.42$, $p = .009$, partial $\eta^2 = .19$, with the reading impairment–only subgroup producing larger clusters than the subgroup with both language and reading impairment, $p = .02$. There were no effects of diagnostic subgroup on embedded cluster ratio, cluster switching, or slope.

In summary, the LD group was less fluent than the ND group. The LD group produced smaller clusters and fewer embedded clusters than the ND group, but the groups did not differ in switch behavior or slope of change over the course of the minute-long interval. There were main effects of category on total correct names, switch behavior, and change over time, but none of these interacted with effects of diagnostic group. Gender was rarely a predictor. The two exceptions were that women repeated cluster types more often than men and that their naming declined more rapidly in the *animal* category than in the *food* category. Variations within the LD group were related to lexical-semantic knowledge, executive control of memory, LD diagnosis, and receipt of accommodation.

Discussion

Semantic Fluency and LD

Students with LD presented with lower semantic fluency than students with ND in both the *animal* and *food* categories. This finding is consistent with findings from related populations including children with specific language impairment (SLI; Henry et al., 2012; Weckerly et al., 2001; but see Marshall, Rowley, Mason, Herman, & Morgan, 2013) and young adults with severe dyslexia (Kinsbourne et al., 1991). On the animal naming task in particular, the proportion of students with LD who performed poorly was more than 2.5 times greater than the proportion of students with ND who performed poorly. That said, effect sizes were small, and overlap between groups was extensive, making it clear that semantic fluency tasks are not useful in the identification of LD in this population. High levels of within-group variability wherein only a minority of affected participants present with vocabulary deficits is also characteristic of developmental language disorders at younger ages (Gray, 2003; Sheng & McGregor, 2010).

Given previous reports (Fournier-Vicente et al., 2008; Hedden et al., 2005; Hughes & Bryan, 2002; Rosen & Engle, 1997; Ruff et al., 1997; Unsworth et al., 2010), we predicted that lexical-semantic knowledge and executive function would be associated with individual differences in fluency between the two participant groups. This prediction held for the group with LD. Individuals who named fewer items also tended to earn lower scores on the PPVT-4 and Trial 1 of the CVLT-II. Thus, semantic fluency tasks may be useful for confirming suspected weaknesses in these areas.

The predicted associations held only in part for the ND group. There was a small but significant correlation between fluency and scores on the PPVT-4. Although there was sufficient range in fluency and CVLT-II scores among the participants with ND to obtain a correlation, there was none. We hypothesize that the semantic fluency task itself is not difficult enough to stress executive function capacities among the ND participants. Were the task to last longer than a minute, it is possible that demands on executive function (and lexical-semantic knowledge) would be greater and the predicted association would obtain.

Lexical Semantics and LD

We analyzed cluster behavior to establish a window into knowledge of relationships between items and the organization of the semantic lexicon. The LD group did cluster their responses, and they did demonstrate embedded clusters; thus, their semantic lexicons are not atypical. In fact, every participant demonstrated semantic clustering. However, their clusters were small and their embeddings infrequent. These quantitative differences between college students with and without LD suggest a less developed semantic lexicon. This conclusion is consistent with the large difference between the vocabulary scores of the two groups. The PPVT-4 and EVT scores of the LD group fell nearly a full standard deviation below the scores of the other college students in this sample. Although these scores were not clinically significant in any traditional sense, they were likely of functional significance in the language-heavy postsecondary context.

Executive Function and LD

Given that switching from cluster to cluster is thought to reflect executive control of memory (Abwender et al., 2001; Troyer et al., 1997; Welsh et al., 1991) and that such conscious control would be increasingly necessary as the minute unfolded (Crowe, 1998), we analyzed cluster switches and change over time to examine the executive control of memory skills of students with LD. The LD group did not differ from the ND group in relative use of cluster switches versus hard switches. This is not to say that verbal short-term memory is irrelevant to the task. Within the LD group, those with lower scores on Trial 1 of the CVLT-II had lower cluster switch to hard switch ratios. Moreover, it does not say that the LD group had strong verbal memory skills; there was a large difference between the LD and ND groups on the CVLT-II that favored the ND group. It does, however, imply that, as a group, the participants with LD demonstrated adequate executive control to get the job done.

Similarity between groups in change over the course of the minute-long naming interval constitutes converging evidence. We were critically interested in whether the LD group were particularly less fluent after the first 15 seconds, the period in which more conscious executive control of memory strategies is required (Crowe, 1998; Luo et al., 2010). Given no significant interaction between diagnostic group and time interval, together with no significant differences between groups in slope, we have no evidence that weakness in executive control of memory interfered with performance. No differences between subgroups of students with LD who did and did not have ADHD—a condition defined by poor executive control—bolstered this conclusion.

This conclusion should be considered in light of Bishop and Norbury (2005). These investigators compared children with SLI, pragmatic language impairment, autism, or no impairments on two measures of ideational fluency. In the first, the children were asked to imagine multiple ways to use an object (e.g., a brick could be used to crush flies). In the second, they were to imagine all of the possible referents for a random line drawing (e.g., two roughly parallel lines could be a road). Note that this task, albeit verbal, does not depend strongly on retrieval of conventional knowledge from the long-term semantic lexicon; instead, success likely depends on executive function, especially cognitive flexibility. With this in mind, it is interesting to note that the children with SLI did not differ from their age-mates on this task (the children with autism and pragmatic language impairment performed more poorly). This finding coincides nicely with the current finding. It seems that the lexical-semantic aspects of the task, not the executive aspects, limited fluency in the LD group.

Future Directions

Task Limitations

By design, verbal fluency tasks are time-limited. Time pressure can highlight differences between diagnostic groups, as subtle deficiencies are more noticeable under demanding conditions. However, the brevity of the task prevents documentation of the full extent of knowledge that any given individual possesses about animals, foods, or any other category of interest. As a follow-up to this study, the depth of lexical-semantic knowledge could be documented by probing exhaustive knowledge of a single category, such as *bird.* This would elicit subordinate knowledge (e.g., *robin sparrow hawk*), and given the lower rate of embedding among the students with LD, probes of subordinate category knowledge could be particularly telling.

Semantic fluency tasks are also limited in category type. The students in this study demonstrated greater fluency in the *food* category than the *animal* category. Perhaps the more direct experience that people have with food in their everyday lives explains this difference. If it is indeed harder to name animals than foods, we would expect more strategic searches of the *animal* category, and this is what is indicated by the higher rate of cluster switches during animal naming than food naming. Also, compared to the *food* category, the *animal* category was more revealing of individual differences: A larger proportion of the LD group performed at a level that could be considered clinically significant during animal naming. Although the utility of probing the *animal* category is clear, other possible categories attested in the literature include types of weather, natural landscape and geographical formations, buildings, and trees (Troyer et al., 1997). Given variations from category to category in size and structure, and differences in the opportunities available for learning them, other categories should be tested to determine those that yield the most useful information about semantic knowledge and organization and to ensure maximum sensitivity to clinical differences.

Sample Limitations

By limiting the sample of participants to young adults in college, we introduced some homogeneity into the data, which is useful for finding patterns but also limits generalization. We have emphasized the relevance of the strategic aspect of executive function to the semantic fluency task, especially as measured by cluster switches, but strategic processes are one of the earlier to develop aspects of executive function (Anderson, Anderson, Northam, Jacobs, & Catroppa, 2001). If we administered a semantic fluency task to much younger children with LD, we might find that executive function deficits limit performance. Moreover, young adults with LD who have not chosen to pursue postsecondary studies or who have been unable to gain admission to postsecondary studies may exhibit more severe deficits than the young adults sampled here, and therefore, the current findings may underestimate the differences in semantic fluency associated with LD. Indeed, in the current sample, severity mattered. Students with lower lexical-semantic knowledge, poorer executive function, and deficits in both reading and spoken language were the poorest performers. Also, students who received classroom accommodations were less fluent than those who do not, which may also be a reflection of severity. Students with accommodations tend to enter postsecondary studies with lower achievement scores than those without accommodations (McGregor et al., 2016).

By including in the sample participants with impairments in spoken language as well as participants with impairments in reading, we introduced some heterogeneity into the data, which complicates interpretation. Heretofore, we have interpreted the difference in performance between participants with reading impairment only and those with both reading plus spoken language impairments as a matter of the severity of their LD. The participants with spoken language impairments only fell between these two subgroups and, statistically, differed from neither. However, this null result should be considered in light of the small subsample. There were nine participants in the language impairment–only subgroup; thus, we had limited power to detect differences between the LI and RI subgroups. Therefore, it could be that the problems of deficient lexical-semantic memory and low semantic fluency are better considered as characteristic of spoken language impairment rather than severe LD. This is highly plausible if those with reading impairment only presented with prototypical cases of dyslexia—a problem of decoding, not comprehension. The correct interpretation awaits a larger sample with documentation of the specific manifestations of the LD subtypes.

Conclusions

To our knowledge, this study is the first to treat cluster switches separately from hard switches when examining the fluency of people with LD. Also, it is the first to consider cluster embedding as a clinically relevant aspect of semantic fluency. Results from this study confirmed the hypothesis that deficits in the lexical-semantic system associated with LD may persist into young adulthood, even among those who have managed their disability well enough to pursue college studies. Lexical-semantic deficits are associated with compromised semantic fluency, and the two problems are more likely among students with more severe disabilities. This study serves to enhance clinicians' understanding of later manifestations of specific language learning disabilities and to motivate interventions aimed at bolstering lexical-semantic knowledge.

Acknowledgments

This research was funded by National Institutes of Health Grant NIH-NIDCD 5R01DC011742 awarded to Karla K. McGregor. We thank Nichole Eden and Tim Arbisi-Kelm for their assistance with data collection and coding and Katherine Gordon for her comments on the manuscript.

References

Abwender, D. A., Swan, J. G., Bowerman, J. T., & Connolly, S. W. (2001). Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods. *Assessment, 8*(3), 323–338.

Aguilar, J. M., & Plante, E. (2014). Learning of grammar-like visual sequences by adults with and without language-learning disabilities. *Journal of Speech, Language, and Hearing Research, 57*(4), 1394–1404.

American Speech-Language-Hearing Association (ASHA). (1990). Guidelines for screening for hearing impairments and middle ear disorders. *ASHA, 32*(Suppl. 2), 17–24.

Americans with Disabilities Act of 1990, Pub. L. No. 101-336, 104 Stat. § 3 (1990).

Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of executive functions through late childhood and adolescence in an Australian sample. *Developmental Neuropsychology, 20*(1), 385–406.

Becker, T. C. & McGregor, K. K. (2016). Learning by listening to lectures is a challenge for college students with developmental language impairment. *Journal of Communication Disorders, 64,* 32–44.

Bental, B., & Tirosh, E. (2007). The relationship between attention, executive functions and reading domain abilities in attention deficit hyperactivity disorder and reading disorder: A comparative study. *Journal of Child Psychology and Psychiatry, 48*(5), 455–463.

Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *International Journal of Bilingual Education and Bilingualism, 10*(3), 210–223.

Bishop, D. V. M., & Norbury, C. F. (2005). Executive functions in children with communication impairments, in relation to autistic symptomatology. *Autism, 9,* 29–43.

Capitani, E., Laiacona, M., & Barbarotto, R. (1999). Gender affects word retrieval of certain categories in semantic fluency tasks. *Cortex, 35*(2), 273–278.

Chen, S. M. (2012). *Lexical organization in Mandarin-speaking children: Insights from the semantic fluency task.* Unpublished master's thesis, University of Iowa, Iowa City.

Crowe, S. F. (1998). Decrease in performance on the verbal fluency test as a function of time: Evaluation in a young healthy sample. *Journal of Clinical and Experimental Neuropsychology, 20*(3), 391–401.

Crowe, S. J., & Prescott, T. J. (2003). Continuity and change in the development of category structure: Insights from the semantic fluency task. *International Journal of Behavioral Development, 27*(5), 467–479.

Delis, D. C., Cullum, C. M., Butters, N., Cairns, P., & Prifitera, A. (1988). Wechsler Memory Scale-revised and California Verbal Learning Test: Convergence and divergence. *The Clinical Neuropsychologist, 2*(2), 188–196.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *Manual for the California Verbal Learning Test–Second Edition (CVLT–II)*. San Antonio, TX: The Psychological Corporation.

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*(5), 1136–1146.

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4 manual*. Bloomington, MN: NCS Pearson.

Ellis Weismer, S., Evans, J., & Hesketh, L. J. (1999). An examination of verbal working memory capacity in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 42*(5), 1249–1260.

Elwér, Å., Keenan, J. M., Olson, R. K., Byrne, B., & Samuelsson, S. (2013). Longitudinal stability and predictors of poor oral comprehenders and poor decoders. *Journal of Experimental Child Psychology, 115*(3), 497–516.

Fidler, L. J., Plante, E., & Vance, R. (2011). Identification of adults with developmental language impairments. *American Journal of Speech-Language Pathology, 20*(1), 2–13.

Fournier-Vicente, S., Larigauderie, P., & Gaonac'h, D. (2008). More dissociations and interactions within central executive functioning: A comprehensive latent variable analysis. *Acta Psychologica, 129,* 32–48.

Gray, S. (2003). Word learning by preschoolers with specific language impairment: What predicts success? *Journal of Speech, Language, and Hearing Research, 47,* 56–67.

Hedden, T., Lautenschlager, G., & Park, D. C. (2005). Contributions of processing ability and knowledge to verbal memory tasks across the adult life-span. *The Quarterly Journal of Experimental Psychology A, 58,* 169–190.

Henry, L. A., Messer, D. J., & Nash, G. (2012). Executive functioning in children with specific language impairment. *Journal of Child Psychology and Psychiatry, 53*(1), 37–45.

Hughes, D. L., & Bryan, J. (2002). Adult age differences in strategy use during verbal fluency performance. *Journal of Clinical and Experimental Neuropsychology, 24,* 642–654.

Hurks, P. P., Schrans, D., Meijs, C., Wassenberg, R., Feron, F. J. M., & Jolles, J. (2010). Developmental changes in semantic verbal fluency: Analyses of word productivity as a function of time, clustering, and switching. *Child Neuropsychology, 16*(4), 366–387.

Im-Bolter, N., Johnson, J., & Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: The role of executive function. *Child Development, 77*(6), 1822–1841.

Individuals with Disabilities Education Improvement Act of 2004 (IDEA), 20 U.S.C. § 1400 (2004).

Johnson, E. S., Humphrey, M., Mellard, D. F., Woods, K., & Swanson, H. L. (2010). Cognitive processing deficits and students with specific learning disabilities: A selective meta-analysis of the literature. *Learning Disability Quarterly, 33*(1), 3–18.

Kail, R., & Leonard, L. B. (1986). Word-finding abilities in language-impaired children [Monograph]. *ASHA Monographs,* (No. 25), 1–39.

Kan, P. F., & Windsor, J. (2010). Word learning in children with primary language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 53*(3), 739–756.

Kapa, L. L., & Plante, E. (2015). Executive function in SLI: Recent advances and future directions. *Current Developmental Disorders Reports, 2*(3), 245–252.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test* [Measurement instrument]. Hoboken, NJ: Wiley.

Kinsbourne, M., Rufo, D. T., Gamzu, E., Palmer, R. L., & Berliner, A. K. (1991). Neuropsychological deficits in adults with dyslexia. *Developmental Medicine & Child Neurology, 33*(9), 763–775.

Leonard, L. B. (2014). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition, 114*(1), 29–41.

Marshall, C. R., Rowley, K., Mason, K., Herman, R., & Morgan, G. (2013). Lexical organization in deaf children who use British Sign Language: Evidence from a semantic fluency task. *Journal of Child Language, 40*(1), 193–220.

Marton, K. (2008). Visuo-spatial processing and executive functions in children with specific language impairment. *International Journal of Language & Communication Disorders, 43*(2), 181–200.

McGregor, K. K. (2014). What a difference a day makes: Change in memory for newly learned word forms over 24 hours. *Journal of Speech, Language, and Hearing Research, 57*(5), 1842–1850.

McGregor, K. K., Arbisi-Kelm, T., & Eden, N. (2016). The encoding of word forms into memory may be challenging for college students with developmental language impairment. *International Journal of Speech-Language Pathology*. Published online April 12, 2016. doi:10.3109/17549507.2016.1159337

McGregor, K. K., Langenfeld, N. R., Oleson, J. J., Van Horne, S. A., Anson, M. K., & Jacobson, W. H. (2016). The university experiences of students with learning disabilities. *Journal of Learning Disabilities Research and Practice, 31,* 90–102.

McGregor, K. K., Newman, R. M., Reilly, R. M., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 45*(5), 998–1014.

McGregor, K. K., Oleson, J., Bahnsen, A., & Duff, D. (2013). Children with developmental language impairment have vocabulary deficits characterized by limited breadth and depth. *International Journal of Language & Communication Disorders, 48*(3), 307–319.

Mervis, C. B., & Crisafi, M. A. (1982). Order of acquisition of subordinate-, basic-, and superordinate-level categories. *Child Development, 53,* 258–266.

Mervis, C. B., & Robinson, B. F. (1999). Methodological issues in cross-syndrome comparisons: Matching procedures, sensitivity (Se), and specificity (Sp). *Monographs of the Society for Research in Child Development, 64*(1), 115–130.

Montgomery, J. W., Magimairaj, B. M., & Finney, M. C. (2010). Working memory and specific language impairment: An update on the relation and perspectives on assessment and treatment. *American Journal of Speech-Language Pathology, 19*(1), 78–94.

Mueller, K. L., & Tomblin, J. B. (2012). Examining the comorbidity of language disorders and ADHD. *Topics in Language Disorders, 32*(3), 228–246.

Poll, G. H., Watkins, H. S., & Miller, C. A. (2014). Lexical decay during online sentence processing in adults with specific language impairment. *Journal of Speech, Language, and Hearing Research, 57*(6), 2253–2260.

Raboutet, C., Sauzéon, H., Corsini, M. M., Rodrigues, J., Langevin, S., & N'Kaoua, B. (2010). Performance on a semantic verbal fluency task across time: Dissociation between clustering, switching, and categorical exploitation processes. *Journal of Clinical and Experimental Neuropsychology, 32*(3), 268–280.

Reilly, S., Bishop, D. V., & Tomblin, B. (2014). Terminological debate over language impairment in children: Forward movement and sticking points. *International Journal of Language & Communication Disorders, 49*(4), 452–462.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General, 126,* 211–227.

Ross, B. H., & Murphy, G. L. (1999). Food for thought: Cross-classification and category organization in a complex real-world domain. *Cognitive Psychology, 38*(4), 495–553.

Ruff, R. M., Light, R. H., Parker, S. B., & Levin, H. S. (1997). The psychological construct of word fluency. *Brain and Language, 57,* 394–405.

Sanford, C., Newman, L., Wagner, M., Cameto, R., Knokey, A. M., & Shaver, D. (2011). *The post-high school outcomes of young adults with disabilities up to 6 years after high school: Key findings from the National Longitudinal Transition Study-2 (NLTS2).* National Center for Special Education Research, NCSER 2011-3004. Menlo Park, CA: SRI International.

Sauzéon, H., Lestage, P., Raboutet, C., N'Kaoua, B., & Claverie, B. (2004). Verbal fluency output in children aged 7–16 as a function of the production criterion: Qualitative analysis of clustering, switching processes, and semantic network exploitation. *Brain and Language, 89*(1), 192–202.

Sheng, L., & McGregor, K. K. (2010). Lexical–semantic organization in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 53*(1), 146–159.

Stothard, S. E., Snowling, M. J., Bishop, D., Chipchase, B. B., & Kaplan, C. A. (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language, and Hearing Research, 41*(2), 407–418.

Troyer, A. K. (2010). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology, 22*(3), 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*(1), 138–146.

Unsworth, N., Spillers, G. J., & Brewer, G. A. (2010). Variation in verbal fluency: A latent variable analysis of clustering, switching, and overall performance. *The Quarterly Journal of Experimental Psychology, 64*(3), 447–466.

Wechsler, D. (1987). *Wechsler Memory Scale–Revised* [Measurement instrument]. New York, NY: The Psychological Corporation.

Weckerly, J., Wulfeck, B., & Reilly, J. (2001). Verbal fluency deficits in children with specific language impairment: Slow rapid naming or slow to name? *Child Neuropsychology, 7*(3), 142–152.

Weiss, E. M., Ragland, J. D., Brensinger, C. M., Bilker, W. B., Deisenhammer, E. A., & Delazer, M. (2006). Gender differences in clustering and switching in verbal fluency tasks. *Journal of the International Neuropsychological Society, 12*(4), 502–509.

Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology, 7*(2), 131–149.

Williams, K. (2007). *Expressive Vocabulary Test*. Circle Pines, MN: American Guidance Service.