McDowd, J., Hoffman, L., Rozek, E., Lyons, K., Pahwa, R., Burns, J. , & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychologia*, 25, 210-225. DOI: 10.1037/a0021531.
Publisher's official version: http://dx.doi.org/10.1037/a0021531. Open Access version: http://kuscholarworks.ku.edu/dspace/.

Paper citation:
McDowd, J., Hoffman, L., Rozek, E., Lyons, K., Pahwa, R., Burns, J. , & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychologia*, 25, 210-225. DOI: 10.1037/a0021531

Keywords: aging and verbal fluency; AD; PD; measuring fluency

Abstract:

- **Objective:** Verbal fluency measures are frequently part of batteries designed to assess executive function, but are also used to assess semantic processing ability or word knowledge. The goal of the present study was to identify the cognitive components underlying fluency performance.
- **Method:** Healthy young and older adults, adults with Parkinson's disease, and adults with Alzheimer's disease performed letter, category, and action fluency tests. Performance was assessed in terms of number of items generated, clustering, and the time course of output. A series of neuropsychological assessments were also administered to index verbal ability, working memory, executive function, and processing speed as correlates of fluency performance.
- **Results:** Findings indicated that regardless of the particular performance measure, young adults performed the best and adults with Alzheimer's disease performed most poorly, with healthy older adults and adults with Parkinson's disease performing at intermediate levels. The exception was the action fluency task, where adults with Parkinson's disease performed most poorly. The time course of fluency performance was characterized in terms of slope and intercept parameters and related to neuropsychological constructs. Speed of processing was found to be the best predictor of performance, rather than the efficiency of executive function or semantic knowledge.
- **Conclusions:** Together, these findings demonstrate that the pattern of fluency performance looks generally the same regardless of how performance is measured. In addition, the primary role of processing speed in performance suggests that the use of fluency tasks as measures of executive function or verbal ability warrants reexamination.

Text of paper begins on next page:

McDowd, J., Hoffman, L., Rozek, E., Lyons, K., Pahwa, R., Burns, J. , & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychologia*, 25, 210-225. DOI: 10.1037/a0021531. Publisher's official version: http://dx.doi.org/10.1037/a0021531. Open Access version: http://kuscholarworks.ku.edu/dspace/.

Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease

Joan McDowd[1], Lesa Hoffman[2], Ellen Rozek[3], Kelly E. Lyons[1], Rajesh Pahwa[1], Jeffrey Burns[1], Susan Kemper[3]

[1]University of Kansas Medical Center, [2]University of Nebraska-Lincoln, [3]Universityof Kansas,

Corresponding author:

Joan McDowd, PhD
Landon Center on Aging
Mail Stop 1005
University of Kansas Medical Center
Kansas City, KS 66160

Voice: 913-588-0646
Fax: 913-588-3179
E-mail: jmcdowd@kumc.edu

McDowd, J., Hoffman, L., Rozek, E., Lyons, K., Pahwa, R., Burns, J. , & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychologia*, 25, 210-225. DOI: 10.1037/a0021531.
Publisher's official version: http://dx.doi.org/10.1037/a0021531. Open Access version: http://kuscholarworks.ku.edu/dspace/.

Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease

Introduction

Executive function (EF) is increasingly recognized as an important factor in the functional status of older adults. This association has been observed in generally healthy community-dwelling older adults (e.g., Carlson et al., 1999; Grigsby et al., 1998; Royall et al, 2004), as well as in people with Alzheimer's Disease (AD) (Swanberg, Tractenberg, Mohs, Thal, & Commings, 2004) and Parkinson's Disease (PD) (Rochester, Hetherington, Jones, Nieuwboer, Willems, Kwakkel, & Van Wegan, 2004). A number of authors have suggested that an individual's executive abilities might be used to anticipate care needs or to plan rehabilitation programs (e.g., Carlson et al., 1999; Kahokehr et al., 2004; Royall et al, 2000). However, a better understanding of measures of EF is necessary in order to realize this potential for clinical significance.

The tasks typically used to assess EF are complex, often involving a number of different cognitive processes. Verbal fluency measures are frequently part of batteries designed to assess EF but are also used to assess semantic processing ability or word knowledge. Since both semantic and executive processes may contribute to performance, it is of interest to isolate the contribution of each to identify the causes underlying performance deficits.

Recent work has begun to develop separate performance measures to isolate semantic and executive performance components (e.g., Beatty et al., 2000; Mayr & Kliegl, 2000; Rosen et al., 2005; Troster et al, 1998; Troyer et al., 1997). Troyer et al. (1997) suggested that "(a) clustering, the production of words within semantic or phonemic categories; and (b) switching, the ability to shift efficiently to a new subcategory" (p. 139) reflect separable semantic and executive processes in fluency performance, respectively. They developed measures of category clustering and category switching for standard fluency tasks as an improvement over simply scoring the total number of words correctly generated. In two experiments, they found that older adults made fewer switches but achieved comparable clustering to the young adults, and that divided attention affected the number of switches but not clustering. They interpreted these findings as support for clustering and switching as "dissociable components" of fluency (p. 143), that age differences in fluency are due to one task component (executive function) and not another (semantic processing) and also that executive function was critical to understanding fluency performance.

An alternative to the Troyer et al method for decomposing fluency performance into semantic and executive components is to examine the time course of responding, as proposed by Mayr and Kliegl (2000). In their task, fluency performance is audio recorded to track the time course of responding. Inter-word response intervals are computed and modeled as a function of retrieval position: $t_n = c + (s * n)$ where t_n is the time between the recall of word n – 1 and word n, c is the constant, or intercept parameter, and s is the slope representing the time increment with every additional word recalled. In this model the constant reflects executive processes and the slope reflects semantic processing ability. That is, executive processes are responsible for initiating and maintaining the retrieval set for each retrieval process, and these task demands are assumed to be constant across retrieval activities. In addition, this model holds that retrieval rate across serial positions (the slope) reflects semantic processing ability. Early in retrieval, high frequency exemplars of the category cue will be retrieved relatively quickly. As the store of these common exemplars is exhausted, additional semantic activation is required to identify additional exemplars. Better semantic ability will lead to quicker identification of these exemplars, producing a shallower slope across serial position.

Mayr and Kliegl (2000) tested this model with category fluency tasks that included a manipulation of semantic retrieval difficulty by varying the frequency and familiarity of the categories (e.g., easy category: animals; hard category: fluids). They also manipulated demands on EF by including both traditional and "switching" versions of fluency tasks. This switching manipulation was motivated by Troyer et al.'s (1997) suggestion that executive processes are particularly relevant in memory retrieval tasks that require switching between semantic clusters. Thus participants were asked to produce category exemplars, either blocked by category in the traditional way (e.g., animal, animal, animal) or alternating between two categories (e.g., animal, tool, animal, tool, animal, tool).

Mayr and Kliegl (2000) hypothesized that the category difficulty manipulation should affect the slope parameter, and the switching manipulations should affect the intercept parameter, and also that age differences in the slope parameter would indicate age differences in semantic processes in fluency tasks, whereas age differences in the intercept parameter would indicate age differences in EF. Mayr and Kliegl found that for the traditional, no-switch task format, no age differences were observed in the slope parameter; further, the difficulty manipulation did affect the slope as expected, but equally so for young and old adults. In contrast, there was a significant age difference in the intercept parameter, hypothesized to reflect EF. "This pattern [supports] the assumption of no age effects in semantic processing (i.e., the slope parameter), but age sensitivity in non-semantic, executive processes" (p. 36) (i.e., the intercept parameter).

In summary, the approach and findings of Mayr and Kliegl and of Troyer et al. illustrate two important points. First, there are significant interpretive problems when task components are poorly understood and task performance could be affected by any of them. Second, separate analysis of task components is very useful in identifying specific deficits. In the case of normal aging, fluency task performance deficits appear to be due more to EF deficits than to semantic memory differences. Importantly, this approach to understanding fluency performance can be extended to advance our understanding of cognitive deficits in age-associated diseases such as AD and PD.

Many of the cognitive deficits observed in PD (for reviews, see Brown & Marsden, 1990; Dubois & Pillon, 1997; Ridenour & Dean, 1999; Troster & Woods, 2003) can be grouped under the heading of "executive function deficits". Taylor and Saint-Cyr (1995) have stated that "the greatest area of difficulty for PD patients unquestionably involves 'executive functions'" (p. 283; see also Cools, Barker, Sahakian, & Robbins, 2001, Rogers et al, 1998; Pollux & Robertson, 2002; Ravizza & Ciranni, 2002; Ravizza & Ivry, 2001; Kensinger, Shearer, Locascio, Growdon, & Corkin; 2003). Fluency deficits are often observed in PD and the contribution of these executive deficits to fluency deficits in PD could be significant. In addition, it has been suggested that individuals with PD are particularly impaired on a novel variant of traditional fluency tests, one requiring them to generate actions such as "things people do" (e.g., Peran, Rascol, Demonet, Celsis, Nespoulous, Dubois, & Cardebat, 2002; Signorini & Volpato, 2005). Signorini and Volpato (2005) indicated that action fluency deficits are consistent with fronto-striatal circuit impairments, suggesting executive dysfunction. Piatt, Fields, Paolo, and Troster (1999) also reported evidence that action fluency deficits in PD were related to poorer performance on tasks such as Trailmaking and the Wisconsin Card Sorting Tasks but not with performance on tests of general cognition or semantic retrieval.

Fluency tasks are also frequently used to assess semantic processing in AD. Understanding these fluency effects is critical as fluency tasks are being proposed as diagnostic tools for AD (Canning, Leach, Stuss, Ngo, & Black, 2004) and it has been suggested that semantic memory impairments may contribute to "the difficulties that AD patients experience in everyday life" (Laatu, Revonsuo, Jaykka, Portin, & Rinne, 2003, p. 82). In addition, impairment in EF is now recognized as a common deficit in AD (e.g., Cummings & Cole, 2002; Swanberg, Tractenberg, Mohs, Thal, & Cummings, 2004; Duke & Kaszniak, 2000; Amieva, Phillips, Della Salla, & Henry, 2004). Again, the relative contribution of semantic and EF deficits to observed fluency deficits is essential to understand what fluency deficits are revealing about cognition and to more precisely define the role of fluency tasks as a diagnostic tool in patients with cognitive decline.

A remaining question is whether the cluster methods proposed by Troyer et al. (1997) or the time-course approach proposed by Mayr and Kliegl (2000) more fully characterizes fluency performance. Troyer and others have used their fluency scoring method with some success in characterizing fluency deficits in AD, PD with and without dementia, and Huntington's disease (e.g., Beatty et al., 2000, Beatty et al., 2002; March & Pattison, 2006; Rich et al., 1999; Testa et al., 1998; Troyer et al., 1998). However, the method's usefulness for discriminating between clinical populations has been questioned (e.g., Epker et al., 1999; Troster et al., 1998). Epker et al. (1999) reported that the measure of total words generated was as good or better than the clustering and switching measures in distinguishing AD, PD, and healthy older adults. In terms of Troyer's two-component model, Abwender et al. (2001) questioned the independence of the switching component, suggesting that the switching measure is only "an index of the *lack* of clustering" (p. 237). Mayr (2002) also raised concerns about the validity of

Troyer's two-component model of fluency. He argued that decreased number of switches in the typical 60-second task period could reflect either a lengthening of switching times or other difficulty with switching (as assumed by Troyer et al., 1997) or a lengthening of response times *within* a cluster. In other words, if retrieval within a cluster is slow, there will be less opportunity to switch because time is limited for task performance. Thus there is ambiguity in the interpretation of the switching component that cannot be disentangled by the Troyer et al. (1997) method (see also Rosen et al., 2005).

We report here a study that compares the utility of traditional scoring methods, the clustering analysis proposed by Troyer et al., and the time course analysis of Mayr and Kliegl. Our goal is to achieve a better understanding of the components of fluency task performance, and to identify the measures that may distinguish patient groups on the basis of semantic and EF processes. We address three primary questions: (1) Do the alternative fluency scoring methods produce different patterns of results across young adults, healthy older adults, adults with AD, and adults with PD? (2) Do individual differences in fluency covary with individual differences in EF and other cognitive abilities? (3) Are group differences in fluency tasks among younger adults, healthy older adults, AD, and PD better characterized as reflecting deficits in semantic processing or in EF?

We address these questions using both experimental manipulations and clinical populations in our study design. To address the question about fluency scoring methods, we report data from a variety of measures of fluency performance, including the number of words correctly generated, the number of clusters generated, the number of words per cluster, and the time course of fluency output. To address the question of whether fluency performance reflects semantic or EF, we include task manipulations designed to separately challenge executive and semantic ability. Following earlier work described above, we used fluency prompts of varying difficulty in order to assess semantic function. We predicted that this manipulation would most negatively affect participants with AD, given the semantic deficit that typically characterizes that population. In addition to fluency prompt difficulty, we compared two task formats: the traditional format and a switching format that requires participants to track and respond alternately to two separate prompts. This alternating format has been assumed to challenge EF, and we expected that it would affect the AD and PD groups more than the other groups in light of the EF deficits that typically characterize those populations. In addition, based on the model proposed by Mayr and Kliegl (2000), we expected the manipulation of semantic difficulty to have an effect on the slope parameter, and the manipulation of task format (traditional, alternating) to have an effect on the intercept parameter in the time course analyses.

The question of whether individual differences in fluency covary with performance on other cognitive abilities was addressed using regression and mixed modeling techniques. We included measures assessing speed of processing, inhibition, working memory, and verbal ability to covary with fluency task parameters and examined their relation to the set of fluency measures. We expected verbal ability and inhibition to be the best predictors of fluency performance, based on their association with semantic and EF, respectively.

Method

Participants

All participants were native English speakers and paid $10 per hour for their time. The thirty-six young adults (18-30 years old, M_{YA} = 21.5, *SD* = 3.1) were enrolled at a Midwestern university. Data from one additional participant was lost due to a technical failure. All young adults were recruited using flyers posted on campus. The healthy community-dwelling older adults (n=30, 65-90 years old, M_{HA} = 72.0, *SD* = 5.4) were recruited from databases of past research participants maintained by the Grayhawk Laboratory at the University of Kansas Medical Center and the Language Across the Lifespan Laboratory at the University of Kansas. Data from three older adults were lost due to technical problems.

Twenty-three community-dwelling older adults (65-90 years old, M_{AD} = 73.8, *SD* = 7.2) with AD were recruited from the participant registry maintained by the Alzheimer and Memory Program at the University of Kansas Medical Center. All participants were evaluated with a standard clinical examination that included neurological and neuropsychological testing and a semi-structured interview with the participant and with a collateral source knowledgeable about the participant, as previously described (Burns, Cronk et al., 2008). Diagnostic criteria for AD require the gradual onset and progression of impairment in memory and in at least one other cognitive and functional domain (McKhann, Drachman et al., 1984). Dementia severity was determined using the Clinical Dementia Rating (Morris, 1993) and all participants had mild (CDR = 1.0) or very mild (CDR 0.5) AD. These methods have a diagnostic accuracy for AD of 93%(Berg, McKeel et al. 1998) and are sensitive to detecting the earliest stages of AD by focusing on intraindividual change rather than comparison with group norms (Storandt, Grant et al., 2006). Fluency testing occurred at the Landon Center on Aging at University of Kansas Medical Center or at a testing site in Lawrence, KS. Data from one participant with AD was lost due to technical problems, another withdrew from testing.

Thirty independent-living individuals with PD (65-90 years old, M_{PD} = 71.9, *SD* = 6.0) were recruited from the University of Kansas Medical Center's Parkinson's Disease Center of Excellence and tested primarily at the Landon Center. . The participants had idiopathic PD and symptoms were mild to moderate. Twenty-one of the 29 participants were Hoehn & Yahr stage 2 (bilateral disease), 7 were stage 3 (bilateral disease with some impairment of balance), and 1 was stage 1 (unilateral disease). Mean UPDRS total score was 34.3 (sd=11.7) and mean subscale scores were 1.4 (1.4) for Mentation, 11.3 (5.6) for ADLs, and 21.5 (6.3) for Motor. Individuals with forms of parkinsonism other than idiopathic disease were excluded from the study. Participants did not have other chronic conditions such as a history of stroke, use of anxiolytics, antidepressants, neuroleptics, sedatives, alcohol abuse, pulmonary disease, and other conditions that may affect speech articulation and speech rate nor did they show any signs of dementia.

Table 1 shows demographic information for the four participant groups. Healthy older adults had completed significantly more years of formal education than the other groups. Scores on the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975), a general assessment of cognitive status including tests of orientation, attention, memory, and language, were lowest among the older adults with AD. All participants were given the Geriatric Depression Scale (GDS; Sheikh & Yesavage, 1986) which is a widely used depression measure. The older adults with PD had the highest scores on the GDS.

Cognitive Tests

A battery of tests of cognitive abilities was administered by a trained research assistant. The test battery included tests of verbal ability, processing speed, working memory, and inhibition. All testing was audio recorded to facilitate later analysis. A *digital ink* software utility developed by the Digital Electronic Core of the Biobehavioral Neurosciences of Communication Disorders Research Center at the University of Kansas was used to administer the tests and record their responses. Table 1 summarizes the results of the cognitive tests. For these and all the comparisons, least significant difference tests were used to compare the groups with alpha = 0.05.

Verbal ability. The Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983) was used to assess verbal ability. Sixty black and white line drawings were presented to the participants to name. Semantic and phonetic cues were provided if the participant was unable to spontaneously name the object. The older adults with AD had significantly lower scores than any of the other participant groups.

Processing speed. The Digit Symbol task (Wechsler, 1958) was used to test speed of processing. Participants matched symbols to numbers for 45 seconds. The number of correct responses was totaled. Young adults performed faster than the 3 groups of older adults. The Letter Comparison Task was also used to measure processing speed. Participants compared pairs of lists of randomly ordered letters to determine if the pairs were the same or different. The pairs were considered different if any of the letters were in a different order or if there were differences in letter identity. Participants were asked to make 30 comparisons as rapidly as possible. The time required to complete the task was recorded, as well as the number correct. The letter comparison rate was calculated using equation 1:

$$\text{Letter Comparison Rate} = \text{time in seconds} / \text{number correct} \quad (1)$$

Young adults were faster on this test and responded more accurately, resulting in faster letter comparison rates.

Inhibition. The Stroop task was used to assess processing speed and inhibition. On the baseline tests, participants identified the color of blocks of colored XXX's (blue, green, red) as quickly as possible for 45 seconds. The number of correct responses was counted. On the Stroop color word test, color words (BLUE, GREEN, RED) were printed in different colors of ink, e.g., "BLUE" printed in green ink. Participants identified the color of the ink, ignoring the words, for 45 seconds, and the number of correct responses was recorded. Inhibition was measured by calculating an interference score with equation 2:

$$\text{Interference} = (\text{blocks of XXX} - \text{color names}) / \text{blocks of XXX} \quad (2)$$

Young adults were fastest on the Stroop XXX and the color word tests. In addition, young adults experienced the least interference, indicating better inhibitory function.

Also used to assess inhibition was the Trail Making test. Trails A was given to measure processing speed as participants connected labeled dots in numerical order, while Trails B required the participant to connect labeled dots in sequential order, alternating between letters and numbers (1-A-2-B-3-C and so forth). This sequence required participants to repeatedly delete a category of responses from the focus of attention. An interference score assessing the participant's ability to carry out these repeated deletions was calculated using equation 3:

$$\text{Interference} = (\text{time in seconds on Trail A} - \text{time in seconds on Trail B}) / \text{time in seconds on Trail A} \quad (3)$$

Older adults with AD were slowest on both Trails A and Trails B although the 3 groups of older adults produced equivalent interference scores on the Trail Making Test.

The Wisconsin Card Sorting Task (WCST) is a classic measure of inhibition and EF. The WCST was administered using a 64 – card computerized version on a Toshiba Tablet PC. A series of displays with squares, crosses, stars, or circles that vary in number or color were presented on the screen and the participant responded to each display based on a rule. The rule was not made explicit by the tester, instead each response was acknowledged as correct or incorrect and the participant had to infer the rule. During the test, after several correct responses a new rule was introduced and the participant must adapt to the change without being told. In addition to correct responses, the total number of perseverative and non-perseverative errors was recorded. Perseverative errors were failures to switch to the new rule, typically occurring during a transition to the new rule. Non-perseverative errors occurred when an individual failed to follow the current rule correctly. Young adults had fewer perseverative errors and fewer non-perseverative errors than any of the older groups. Healthy older adults committed fewer perseverative errors than either group of impaired older adults. The healthy older adults and older adults with PD had similar numbers of non-perseverative errors but both made fewer non-perseverative errors than the older adults with AD.

Working memory. To assess working memory, the Digits Forward and Digits Backward tests were administered (Wechsler, 1958). Participants listened to a string of numbers between 2 and 10 digits in length which they were to repeat in the same order (Digits Forward) or in reverse order (Digits Backwards). The Operation Span (OSPAN) test was also given as a measure of working memory. For this task, the participant read an arithmetic equation out loud, responded whether the equation was correct or not, then read a word printed beside the equation on the page. The participant repeated this for a set of 1 to 5 equations; at the end of each set of equations, the participant verbally recalled each of the words from the set. Older adults with AD performed poorly on these tests, especially the OSPAN test.

The cognitive battery was designed to assess four individual differences that have been assumed to contribute to performance on verbal fluency tests: verbal ability, working memory, inhibition, and processing speed. A principal components analysis with varimax rotation was used to extract for 4 factors. Education and the Boston Naming Test defined a verbal ability component, accounting for 58.66% of the variance between participants on these measures; the span scores yielded a working memory component, accounting for 62.69% of the variance; an inhibition component was obtained from the Stroop and Trail Making Interference Scores, and the number of correct responses on the WCST, accounting for 55.88% of the variance; and a processing speed component was derived from the Digit Symbol, Letter Comparison, Trails A, and Stroop XXX measures, accounting for 81.58% of the variance. Verbal ability was weakly correlated with the other components, rs < .10; processing speed was moderately correlated with inhibition, $r = .52$, and working memory, $r = .36$; and working memory and inhibition were moderately correlated, $r = .46$.

Verbal Fluency Tests

Materials. Verbal fluency tests were interspersed among the other cognitive tasks described above. Three types of verbal fluency tests were administered; letter (phonetic) fluency (using letters S/L/M/P), semantic fluency (Easy categories – birds, clothes, body parts, colors; Hard categories – insects, fabrics, fluids, writing utensils), and action fluency (things people do; ways you can move; ways you can talk; things you can do to an egg). Prompts for these tests were selected based on pilot testing with a group of 10 young adults and 10 healthy older adults. To select the letter prompts, each participant in the pilot study was given a list of 12 consonants and asked to generate as many words as possible beginning with each letter, excluding proper names. Participants were allowed 15 minutes to complete the task. The final set of letter prompts was selected such that each participant generated 40 to 50 exemplars of each letter (mean = 43) by both young and older adults. Rejected letter prompts resulted in 25 or fewer exemplars by each group. The category prompts were selected from a wider set of potential prompts based initially on pilot testing with the group of 10 young and 10 older adults; the participants were allowed 15 minutes to generate exemplars of 30 different category prompts. Four prompts that resulted in 50 to 60 exemplars (mean = 56) from each group of participants were selected as easy categories; hard categories elicited 15 to 30 (mean = 26) from each group of participants. The easy / hard distinction was then validated against two sets of word association norms. Category norms provided by Van Overschelde, Rawson, and Dunlosky (2004) provide category norms including a measure of "category potency": the average number of responses given for each category by participants within 30 s. Easy categories resulted in 6 or more responses, on average (mean = 8.6); hard categories resulted in 6 or fewer responses (mean = 4.3). In addition, easy and hard categories differed in terms of the number of exemplars listed for each prompt in the South Florida word association norms (Nelson, McEvoy, & Schreiber, 1998). Easy categories had 50 or more exemplars (mean = 83); hard categories had 30 or fewer (mean = 29). Action fluency prompts were developed following Piatt, Fields, Paolo, and Troster (1999) who used a single item "things that people do." Like the letter and category prompts, the action prompts were chosen from a wider set of potential actions based on pilot testing with the group of 10 young and 10 healthy older adults. Each was given a list of 30 actions and asked to generate exemplars. Participants were allowed 15 minutes to complete the task. The final set of action prompts was selected such that each participant generated 20 to 30 exemplars of each action (mean = 28). Rejected action prompts resulted in 15 or fewer exemplars.

Procedure. For the traditional test format, one prompt was provided and the participant was given 3 minutes to ensure an adequate number of responses for the clustering analyses. In the alternating format, participants were given 2 prompts and 6 minutes to respond. [1]

On each test, the participants were shown the prompt(s), e.g. "words that begin with S" written out on a piece of paper, and asked to generate as many words as they could think of that would

[1] Traditionally, verbal fluency is assessed with a shorter response interval, often 30 s or 1 min. On a series of pilot tests, few clusters were observed during a 1 min response interval so a longer response interval was used. To investigate how the length of the response interval affected the outcomes, responses during the 1st and 3rd min of the letter and hard categories traditional tests (1st 2 and last 2 min of the letter and hard categories alternating tests) were compared separately for all 4 groups of participants. On each test, approximately 70% of all responses occurred during the 1st min on the traditional tests and during the 1st 2 min of the alternating tests. The proportion of correct responses produced was consistent as was the average cluster size although the total number of clusters declined from the 1st to the last min (1st 2 to last 2 min of alternating tests). The number of intrusions and perseverations tended to increase from the 1st to the 3rd min (1st 2 to last 2 min of the alternating tests) for all participants although the proportion of intrusions and perseverations was consistent. Thus, all 4 groups of participants tended to respond to the fluency tests with a 'burst' of predominately correct responses, and their output of correct responses gradually diminished while intrusions and preservations gradually increased. The time course analysis provides a more detailed look at this response pattern.

matched the prompt(s). For the alternating tests, the participants were shown both prompts and were instructed to respond to the first prompt, then the second prompt, and continue back and forth. The page showing the prompt(s) remained available for the participant to look at the entire time. Two traditional fluency tests and one alternating fluency test was administered for each type of fluency, e.g. letters, easy categories, hard categories, and actions. Prompts were counterbalanced across participants, e.g. one participant responded to S and L prompts in the traditional format and M and P prompts in the alternating format; while the next participant responded to M and P prompts in the traditional format and S and L prompts in the alternating format.

Fluency Responses. All responses were digitally recorded for later transcription. The transcripts were coded for 4 types of responses: correct responses, perseverations, intrusions, and clusters. In addition, cluster size and cluster switches were calculated. Meta-comments were counted as well but were not included in the analyses. Meta-comments are comments made by the participant about their own performance or attempting to avoid responding such as "I am not doing so well," "This is really difficult," "Are these responses ok?" or "I don't know." Correct responses were required to meet the target rule. Perseverations were correct responses that were repeated; e.g. if the participant responded, "robin, bluebird, cardinal, robin," the second occurrence of "robin" was scored as a perseveration. Intrusions were incorrect responses that were not members of the category; if the participant responded, "hand, foot, shirt, head," shirt was scored as an intrusion. Clusters were a group of words generated successively that formed a subgroup of the category. For letter fluency, clusters were identified as words that rhymed (mint, meant), differed by vowel (mat, mitt), shared an initial consonant cluster (straight, stop), or were homonyms (pair, pear, pare) (Troyer, Moscovitch, & Winocur, 1997). Clusters for the semantic and action categories were defined for each category. For example, subcategories of colors included rainbow colors, pastels, colors preceded by descriptive words, and shades of the same color. For the action category, subcategories for "things you can do to an egg" include ways to cook an egg, ways to destroy an egg, and ways to decorate an egg. Three measures of clustering were obtained: the number of clusters, cluster size, and the number of switches between clusters. Cluster size was the total number of responses within a cluster minus one. Cluster switching was the number of transitions that occurred between clusters including single word clusters. On the alternating versions, clusters and cluster switching were counted for each prompt separately and the totals for each prompt were summed for a single score.

All audio recorded responses were transcribed and scored by a single coder; 15% (5 per group) of the transcripts were randomly selected and verified by a second coder. Agreement was quite high for the number of correct responses and for identification of perseverations and intrusions (all Cronbach's alpha > .95). Agreement was also high for the identification of clusters, switches, and the measure of cluster size (Cronbach's alpha > .90).

Fluency time course. To examine the time course of responding, inter-word response times (IWTs) were determined by inspection of the speech wave form using Boersma and Weenink's PRAAT sound editing program (version 4.6.02 for Windows; downloaded from www.praat.org). Only the first 10 responses by each participant were timed, yielding 9 IWTs. A coder listened to the recording and marked the onset and offset of each response. A second coder timed 10% of responses; agreement was high, averaging +/- 35 ms. Only IWTs for correct responses were included in the time course analyses although incorrect responses were used to determine the retrieval position and IWTs of correct responses. A time limit of 30 sec was imposed on all IWTs to reduce the impact of extreme values; IWTs for correct responses greater than 30 sec were trimmed to 30 sec. Many participants were unable to generate 10 correct responses in the time allowed, primarily in the Hard category, alternating fluency condition. In that condition, 75% of all participants generated 8 correct responses. In the Action alternating fluency condition, 80% of all participants generated 10 correct responses. All participants produced at least 10 correct responses in all of the other task conditions. Analyses included IWTs for up to 10 correct responses in all task conditions.

Results

Verbal Fluency Responses

The initial series of analyses examined verbal fluency performance using traditional counts of correct responses, perseverations, and intrusions, as well as three measures of clustering: the total number of clusters, average cluster size, and the number of switches between clusters. These measures were obtained from the entire 3 minute response interval for the traditional tests and the 6 minute interval for the alternating tests.

Repeated measures ANOVAs were used to compare the four groups (Young Adults, Healthy Older Adults, Older Adults with PD, Older Adults with AD) on each fluency test (Letters, Easy, Hard, Actions). Format (traditional versus alternating) was a within-subject factor. The ANOVAs are summarized in Table 2 (detailed summary data are available from the authors by request).

The group differences were very similar across fluency measures and types: in general, young adults produced the most correct responses, fewest perseverations, fewest intrusions, most clusters, largest clusters, and most switches and the older adults with AD produced the fewest correct responses, most perseverations, most intrusions, and fewest clusters, smallest clusters, and least switches. Healthy older adults and older adults with PD tended to produce intermediate numbers of correct responses, perseverations, and clusters. The traditional tests produced more correct responses than the alternating tests and this difference was constant across groups. Although perseverations were not affected by format, intrusions were affected: there were fewer intrusions produced on the alternating tests than on the traditional tests and this difference was constant across groups. Intrusions, however, were rare. More clusters and larger clusters resulted from the traditional tests than the alternating tests. These differences were constant across groups, with 2 exceptions for cluster size: on the letter and hard categories tests, the advantage of the traditional format over the alternating format was reduced for the older adults with AD. The format advantage was also consistent across groups for the number of switches, with the exception of the easy category test. In this case, the advantage of traditional over alternating format was also reduced for the adults with AD.

- *Relationship of fluency to other cognitive domains.* To examine how individual differences in processing speed, verbal ability, working memory, and inhibition affected performance on the verbal fluency tests, a series of stepwise regression models was compared using forward selection to control the order of entry of the component scores in the models. For young adults, these models were nonsignificant for both traditional and alternating fluency tests, likely reflecting the limited range of scores among the young adults on these tests of processing speed, working memory, inhibition, and verbal ability. These nonsignificant models are not reported.

- For older adults, processing speed was the best predictor in Step 1 of the number of correct responses on all traditional fluency tests. Inhibition also contributed to traditional number of correct responses, accounting for additional variance in Step 2 for letters, easy categories, and actions. Processing speed was the best predictor of the number of clusters on the letter, easy categories, and action fluency tests and low verbal ability was the best predictor of perseverations on easy and hard categories and actions (see Table 3).
- On the alternating fluency tests, processing speed and inhibition again were independent predictors of correct responses by older adults on letters and easy and hard categories, and processing speed was the sole predictor of correct responses on actions. Processing speed and inhibition were independent predictors of the number of clusters on both alternating category fluency test whereas processing speed and, somewhat surprisingly, working memory were the only significant predictors of the number of clusters on alternating letter and action fluency, respectively. Low verbal ability also predicted perseverations on the alternating easy category and action fluency tests as well as intrusions on the alternating hard category test (see Table 4).
- *Time course of responding.* The next set of analyses allowed us to examine two additional parameters of fluency performance: the slope and intercept describing the time course of serial verbal output. Mayr and Kliegl (2000) hypothesized that the slope represents semantic processing ability and the intercept represents EF. We examined group and task effects on these parameters, along with their neuropsychological underpinnings. Our analyses examined the time course of fluency responding using mixed effects regression models in SAS PROC MIXED for each of the 8 verbal fluency tests and assessed the extent to which these trajectories differed by age, cognitive status, test format, or test difficulty. Maximum likelihood (ML) was used in estimating model parameters and to compare the fit of nested models; denominator degrees of freedom were estimated using the Satterthwaite method (see Snijders & Bosker, 1999, for further information on multilevel or mixed effect models). Although these models are commonly used for longitudinal data, in our case, rather than estimating intercepts and slopes that describe changes over occasions, we are describing changes over the order of response in the fluency task. Specifically, for the alternating versions of the fluency tasks, IWTs were averaged for each response across the 2 series (e.g., the time between the first and second responses for words that begin with S and the time between the first and second responses for words that begin with L were averaged, the time between the second and third responses to S and the second and third responses to L were averaged, and so forth). IWT

McDowd, J., Hoffman, L., Rozek, E., Lyons, K., Pahwa, R., Burns, J. , & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychologia*, 25, 210-225. DOI: 10.1037/a0021531.
Publisher's official version: http://dx.doi.org/10.1037/a0021531. Open Access version: http://kuscholarworks.ku.edu/dspace/.

trajectories using model-estimated intercepts and slopes are shown in Figures 1 – 4 for both the traditional and alternating task versions.

- For each verbal fluency test, the level-1 or within-person model predicted IWTs as a function of retrieval position, and was centered at 1 so that the intercept represented the grand mean latency from the first correct response to the second correct response. A linear slope for the average rate of increase in IWT for successive correct responses was also included, plus a within-person residual for the remaining deviation of each observation from the best-fit line. Although additional nonlinear slope parameters were also examined, their effects were small and nonsignificant, and thus they were not retained in the models.
- The level-2 or between-person model was then augmented sequentially. Model 1 allowed intercepts to vary randomly over persons, but not the linear slopes, which served as a baseline model. Model 2 added a variance across persons for the individual random linear slopes, as well as a covariance among the intercepts and slopes, in order to examine whether the effect of response order on IWT did indeed vary over persons. Model 3 added three fixed effects to distinguish the intercepts and three fixed effects to distinguish the linear slopes across the four groups, with younger adults as the reference group, in order to examine group differences in the order effects in the fluency task. Finally, Model 4 included heterogeneous variance components (random intercepts, random slopes, residuals) as a function of age (young adults vs. all older adults) in order to control for the greater heterogeneity likely to be present among the older adults. Model 4 provided the best fit for all 8 fluency tests; estimates of intercepts and slopes from this Model for each of the 4 groups (young adults, healthy older adults, older adults with AD, and older adults with PD) on each of the 8 fluency tests were used to derive to the trajectories shown in Figures 1 to 4.

Figures 1 – 4 indicate that the estimated intercepts representing mean IWT between the first and second correct responses were similar for young and healthy older adults. The intercepts for older adults with PD tended to be somewhat higher than those for healthy older adults, and the highest intercepts (or longest times between the first and second response) were obtained from older adults with AD. The estimated slopes representing the linear rate of increase in IWT across serial output position varied by group as well. The slopes for young and healthy older adults tended to be similar, the slopes for older adults with PD tended to parallel those for healthy older adults, and the slopes for older adults with AD tended to be significantly steeper. An important exception to these general trends concerns action fluency. As indicated in Figure 4, both the traditional and alternating action fluency tests were particularly difficult for the older adults with PD, resulting in exceptionally steep slopes.

Based on Mayr and Kliegl's (2000) suggestion, we predicted that our manipulation of test format (traditional vs. alternating) would influence the intercept of IWT trajectories. We found some evidence in support of this hypothesis that the additional demands placed on EF by the alternating format would affect IWT intercepts. For example, Table 9 suggests that the traditional and alternating fluency tests yielded similar slopes but different intercepts. To examine this possibility, traditional and alternating fluency tests were directly compared using a multivariate version of Model 4 that included contrasts for each parameter by test format. As shown in Table 5, higher intercepts for all 4 groups of participants were found for the alternating easy category and action fluency tests as compared to the traditional tests. The intercepts for letters and hard categories did not vary with test format, except for letter fluency in healthy older adults and action fluency in older adults with AD. However, the alternating test also produced higher slopes for hard categories and actions, as well as steeper slopes for easy and hard categories in the younger adults. Thus we conclude that the alternating format does result in higher intercepts than the traditional format although it also affects slopes, especially for the more difficult hard category and action fluency tasks.

Similarly, Mayr and Kliegl (2000) suggested that category difficulty would affect the slope of IWT trajectories by affecting semantic memory search processes. To explore this possibility, multivariate models were also used to compare the easy and hard category tasks. As shown in Table 6, slopes were steeper for the hard categories for all 4 groups. However, intercepts also were higher for the two groups of impaired older adults, again providing only partial support for the Mayr and Kliegl interpretation of these parameters.

The comparison of healthy older adults, and older adults with AD and PD was also designed to examine Mayr and Kliegl's hypothesis that IWT intercepts reflect EF and IWT slopes reflect semantic function. We expected intercepts to be higher for older adults with PD and AD due to impairments of EF, and IWT slopes to be higher for older adults with AD due to deficits in semantic function. We found that intercepts were similar for young and healthy older adults, and that intercepts for older adults with PD tended to be somewhat higher than those for healthy older adults. The highest intercepts (or slowest initial responses) were obtained from older adults with AD. This finding of greater intercepts among the groups assumed to have greater executive impairment is consistent with the Mayr and Kliegl hypothesis.

The estimated slopes representing the linear rate of increase in IWT across trials varied by group as well. In this case, the slopes for young and healthy older adults tended to be similar, the slopes for older adults with PD tended to parallel those for healthy older adults, but the slopes for older adults with AD tended to be significantly steeper. Because people with AD are characterized by significant semantic deficit, these group differences in slope support Mayr and Kliegl's interpretation of the slopes as reflecting semantic function. An important exception to these general trends concerns action fluency. As indicated in Figure 4, both the traditional and alternating action fluency tests were particularly difficult for the older adults with PD, resulting in exceptionally steep slopes. These data indicate a specific deficit in verb generation among people with PD, consistent with the findings of Peran et al. (2002) and Signorini and Volpato (2005).

Another approach to testing the model of IWT slope and intercept parameters was to examine the relation among these parameters and performance on the cognitive domains assessed by the cognitive battery administered here. We predicted that individual differences in verbal ability and processing speed would affect IWT slopes by influencing the rate of retrieval of information from semantic memory, and that individual differences in EF components of working memory and inhibition would affect IWT intercepts, especially for alternating fluency tests. To test these predictions, a series of conditional models was specified to investigate the individual and cumulative between-person individual differences. These models included response order as well as additional predictors for the effects of verbal ability, processing speed, inhibition, and working memory on the slopes and intercepts. Parallel to the regression analyses (above), the conditional models were estimated for the 3 groups of older participants only, as summarized in Table 7. Because of the interest in individual differences rather than group type, group effects were not included.

Intercepts were generally associated with processing speed and inhibition, and to a lesser extent working memory, such that slower participants, those with reduced inhibition, and those with less working memory had higher intercepts. Likewise, slopes were generally associated with processing speed and inhibition, and to a lesser extent working memory, such that slower participants, those with reduced inhibition, and those with less working memory had steeper slopes. Verbal ability was not a robust predictor of slopes for either the traditional or alternating fluency tests. An additional model included all 4 between-person predictors; it suggests that each predictor contributes independent variance to modeling both the slopes and the intercepts, with the exception of the intercept for traditional letter fluency where inhibition and working memory appear to share variance. Again, these data do not support a clean dissociation between the cognitive underpinnings of slope and intercept parameters.

Tables 8 and 9 provide an indication of the improvement in model fit by the inclusion of the individual between-person predictors as well as all 4 predictors. In general, the between-person predictors improved model fit (Table 8), the greatest improvement in model fit resulted from the inclusion of the processing speed, inhibition, and/or working memory predictors. Including all 4 predictors resulted in further improvement in model fit, suggesting there is considerable unique variance contributed by the individual predictors. The pseudo-R^2s in Table 9 are the square of the correlation between the predicted values from the fixed effects and the actual outcomes, and can serve as a general indicator of total reduction in outcome variance by the model fixed effects (Snijders & Bosker, 1999).

Discussion

The study reported here was designed to better understand the cognitive components of verbal fluency tasks, which are among the mostly commonly used neuropsychological measures of cognitive function. We compared alternate measures of fluency across task types and participant groups in order to address three primary research questions. Our findings relative to these questions are discussed below.

*Are age and group differences in verbal fluency consistent across alternative scoring methods? O*verall, very similar patterns of age and group differences were observed, regardless of the specific scoring method used to assess verbal fluency performance. In particular, clustering methods produced the same pattern of outcomes as did the traditional method of examining correct responses, perseverations, and intrusions. Although adults with AD showed somewhat less clustering than the other groups, clustering overall was quite infrequent. None of the groups appeared to use the sort of systematic retrieval strategies that would have produced large clusters, obviating group discriminations on the basis of the number of clusters, the size of clusters, or cluster switching. Similarly, the pattern of results obtained with the alternating test format mirrored those obtained from the traditional test format. Although Troyer et al. (1997) argued that clustering and switching were "dissociable components" of fluency, the two measures produced very similar patterns of results, and the number of clusters produced was most strongly associated with speed of processing measures, rather than executive function or semantic memory.

- We conclude that alternative methods for scoring verbal fluency performance and alternative formats for assessing verbal fluency yield very similar outcomes. Our findings suggest that any controversy about the "best" way to measure fluency performance may be unfounded – our data indicate that alternative scoring methods and test formats are equally sensitive to age group differences in verbal fluency as well as to the effects of AD and PD on verbal fluency.
- Analyses of the slope and intercept parameters in the time course analyses revealed a somewhat different pattern of results. In the case of the slope parameter, the performance of adults with AD distinguished them from the other three groups, who performed comparably to one another. Analyses of the intercept parameter showed a different pattern – those with AD produced intercepts higher than adults with PD, and those with PD produced intercepts higher than healthy old or young adults, who did not differ from one another. Thus the slope parameter appears to be particularly sensitive to the deficits associated with AD, whereas the intercept may be sensitive to neurodegenerative disorders, as in these data the intercepts of those with either AD or PD were elevated relative to healthy older or young adults. Next we turn to the question of whether these increased slopes and intercepts can be attributed to deficits in semantic memory and EF, respectively.

Do individual differences in fluency covary with individual differences in executive function and other cognitive abilities?

- Our findings indicate that the same underlying cognitive processes account for verbal fluency performance whether fluency is assessed by the number of words correctly generated, by clustering, or by slopes and intercepts of inter-word response times. By and large, processing speed was the most consistent predictor of overall performance as well as the time-course of responding. Inhibition also contributed to overall verbal fluency performance as well as the time-course of responding. Overall, those generating the most correct responses were also the fastest and had good inhibition. Interestingly, and contrary to expectations, verbal ability did not predict overall performance on any of the fluency tests or the time-course of responding. Fluency deficits in the slope parameter evidenced among people with AD may be due more to slowed responding and poor inhibition than to semantic processing deficits. Also contrary to expectations, individual differences in inhibition contributed to the time-course of responding on both traditional and alternating fluency tasks, suggesting that both test formats affect EF.
- We do note that action fluency does appear to be differentially sensitive to the effects of PD, as suggested by Signorini and Volpato (2006) and Peran et al. (2002). Further, the

effect for processing speed on action fluency was exacerbated among those with PD. Although action fluency has been taken to reflect EF (e.g., Piatt, Fields, Paolo, and Troster, 1999; Signorini & Volpato, 2006), the present results indicate that it may be more closely aligned with speed of processing. The Piatt et al. study was designed to validate action fluency as an EF measure, but they did not include speed of processing measures among those they administered. Future work should include speed measures to most accurately identify task components.

Are group differences in fluency tasks among younger adults, healthy older adults, adults with AD, and adults with PD better characterized as reflecting deficits in semantic processing or in executive functions?

The answer to this question appears to be: NEITHER. In our data, performance differences primarily reflected differences in processing speed. Inhibitory function did play a secondary role for many of the tasks and measures, but verbal ability had only a relatively small influence in a few tasks and measures. Thus for these groups, fluency tasks seem chiefly to be measures of processing speed.

None of the neuropsychological measures predicted fluency performance among young adults. This outcome is likely a result of relatively little variance to be explained in the homogeneous sample of young adults tested for the present study. Thus it remains for future research to determine whether the same pattern of relationships (using more sensitive measures) will hold for young adults as was observed for older adults in the present study.

Summary and Conclusions

Measures of EF are often complex, and plagued by what might be called the "task indeterminacy" problem. That is, it is often unclear exactly what EF tasks measure. Verbal fluency is one such measure – besides being used as an index of EF, it has been used to assess semantic memory and verbal ability. In addition, a variety of new measures and task versions (e.g., clustering, switching between clusters, alternating formats) have been developed, with proponents claiming that each provides a more sensitive measure of performance or best reveals the deficits characterizing specific populations. However, these alternatives have not been systematically and consistently examined in the same sample of both healthy adults and adults with neurological disorders. Our findings indicate that generally the same pattern of age and group differences are observed, regardless of the specific verbal performance measures or format used.

The present findings are also surprising in terms of what fluency tasks actually measure. Our analyses indicate that variations in fluency performance are predominantly accounted for by individual and group differences in processing speed and in inhibitory processes. Somewhat surprisingly, individual and group differences in verbal ability played a very minor role in verbal fluency performance once both processing speed and inhibition were considered. Given this finding, future studies involving fluency tasks should also include measures of processing speed in order to more accurately interpret group or task differences in performance.

And finally, the present study illustrates the challenges of assessing EF with neuropsychological assessments involving multiple (unexpected) components. We addressed this issue in the context of the verbal fluency task, using mixed effects regression techniques to identify fluency task components. Another approach has been taken by Miyake (Miyake, Friedman, Emerson, Witzki, Howerter, & Wager, 2000), Salthouse (Salthouse, Atkinson, & Berish, 2003), and Hull (Hull, Martin, Beier, Lane, & Hamilton, 2008) to use structural equation modeling to identify the subcomponents of EF. Our measures of EF most closely map onto the inhibition and shifting parameters of these models. Other EF components such as updating and dual tasking may also contribute to verbal fluency performance. Clearly, additional work along these lines is necessary in order to identify the task components responsible for performance deficits, both for neuropsychological test performance and everyday functional tasks.

References

Abwender DA, Swan JG, Bowerman JT, Connolly SW. (2001) Qualitative analysis of verbal fluency output: review and comparison of several scoring methods. *Assessment, 8*, 323-38.

Amieva, H., Phillips, L., Della Sala, S., & Henry, J. D. (2004). Inhibitory functioning in Alzheimer's disease. *Brain, 127,* 949-964.

Beatty, W. W., Salmon, D. P., Testa, J. A., Hanisch, C., & Troster, A. I. (2000). Monitoring the changing status of semantic memory in Alzheimer's disease: an evaluation of several process measures. *Neuropsychology, 7*, 94-111.

Beatty, W. W., Salmon, D. P., Troster, A. I., & Tivis, R. D. (2002). Do primary and supplementary measures of semantic memory predict cognitive decline by patients with Alzheimer's disease? *Aging, Neuropsychology and Cognition, 9*, 1-10.

Berg, L., D. W. McKeel, Jr., et al. (1998). "Clinicopathologic Studies in Cognitively Healthy Aging and Alzheimer Disease: Relation of Histologic Markers to Dementia Severity, Age, Sex, and Apolipoprotein E Genotype." *Archives of Neurology, 55*, 326-335.

Burns, J. M., B. B. Cronk, et al. (2008). Cardiorespiratory fitness and brain atrophy in early Alzheimer disease. *Neurology, 71***:** 210-216.

Canning, S. J., Leach, L., Stuss, D., Ngo, L., Black, S. E. (2004). Diagnostic utility of abbreviated fluency measures in Alzheimer disease and vascular dementia. *Neurology, 62,* 556-562.

Carlson, M., Fried, L., Xue, Q., Bandeen-Roche, K., Zeger, S., & Brandt, J. (1999). Association between executive attention and physical functioning performance in community-dwelling older women. *Journal of Gerontology: Social Sciences, 54B*, 262-270.

Cools, R., Barder, R.A., Sahakian, B.J., & Robbins, T.W. (2001). Mechanisms of cognitive set flexibility in Parkinson's disease. *Brain, 124*, 2503-2512.

Cummings, J. L., & Cole, G. (2002). Alzheimer disease. *Journal of American Medical Association, 287*, 2335-2338.

Dubois, B., & Pillon, B. (1997). Cognitive deficit in Parkinson's disease. *Journal of Neurology, 244*, 2-8.

Duke, L. M., & Kaszniak, A. W. (2000). Executive control functions in degenerative dementias: A comparative review. *Neuropsychology Review, 10*, 75-93.

Epker, M. O., Lacritz, L. H., & Cullum, C. M. (1999). Comparative analysis of qualitative verbal fluency performance in normal elderly and demented populations. *Journal of Clinical and Experimental Neuropsychology, 21*, 425-434.

Folstein MF, Folstein SE, McHugh PR. Mini-Mental State: A practical guide for grading the cognitive state of patients for the clinician. (1975). *J Psychiatr Res,* ***12***,189–198

Grigsby, J., Kaye, K., Baxter, J., Shetterly, S., & Hamman, R. (1998). Executive cognitive abilities and functional status among community-dwelling older persons in the San Luis Valley health and aging study. *Journal of the American Geriatrics Society, 46,* 590-596.

Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. (1982). A new clinical scale for the staging of dementia. *British Journal of Psychiatry, 140*, 566-572.

Hull, R., Martin, R., Beier, M., Lane, D., & Hamilton, A.C. (2008). Executive function in older adults: a structural equation modeling approach. *Neuropsychology, 22,* 508-522.

Kahokehr, A., Siegert, R. J., & Weatherall, M. (2004). The frequency of executive cognitive impairment in elderly rehabilitation inpatients. *Journal of Geriatric Psychiatry and Neurology, 17*, 68-72.

Kaplan, E., Goodglass, H., Weintraub, S. (1983). *The Boston Naming Test*. Philadelphia: Lea & Febiger.

Kensinger, E. A., Shearer, D. K., Locascio, J. J., Growdon, J. H., & Corkin, S. (2003). Working memory in mild Alzheimer's Disease and early Parkinson's Disease. *Neuropsychology*, *17*, 230-239.

Laatu, S., Revonsuo, A., Jaykka, H., Portin, R., & Rinne, J. O. (2003). Visual object recognition in early Alzheimer's disease: Deficits in semantic processing. *Acta Neurologica Scandinavica, 108,* 82-89.

March, E. G., & Pattison, P. (2006). Semantic verbal fluency in Alzheimer's disease: approaches beyond the Traditional Scoring System. *Journal of Clinical and Experimental Neuropsychology, 28*, 549-566.

Mayr, U. (2002). On the dissociation between clustering and switching in verbal fluency: comment on Troyer, Moscovitch, Winocur, Alexander and Stuss. *Neuropsychologia, 40*, 562-566.

Mayr, U. & Kliegl, R. (2000). Complex semantic processing in old age: Does it stay or does it go? *Psychology & Aging. 15*, 29-43.

McKhann, G., D. Drachman, et al. (1984). "Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease." *Neurology, 34*: 939-944.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T.D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49-100.

Morris, J. C. (1993). "The Clinical Dementia Rating (CDR): current version and scoring rules." *Neurology, 43*(11): 2412b-2414.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/.

Peran, P., Rascol, O., Demonet, J-F., Celsis, P., Nespoulous, J-L., Dubois, B., & Cardebat, D. (2003). Deficit of verb generation in nondemented patients with Parkinson's disease. *Movement Disorders, 18*, 150-156.

Piatt, A. L., Fields, J. A., Paolo, A., & Troster, A. I. (1999). Action (Verb Naming) fluency as a unique executive function measure: Convergent and divergent evidence of validity. *Neuropsychologia, 37*, 1499–1503.

Pollux, P. M. J., & Robertson, C. (2002). Reduced task-set inertia in Parkinson's disease. *Journal of Clinical and Experimental Neuropsychology, 24*, 1046-1056.

Ravizza, S. M., & Ivry, R. B. (2001). Comparison of the basal ganglia and cerebellum in shifting attention. *Journal of Cognitive Neuroscience, 13*, 285-297.

Ravizza, S. M., & Ciranni, M. A. (2002). Contributions of the prefrontal cortex and basal ganglia to set shifting. *Journal of Cognitive Neuroscience, 14*, 472-483.

Rich, J. B., Troyer, A. K., Bylsma, F. W., & Brandt, J. (1999). Longitudinal analysis of phonemic clustering and switching during word-list generation in Huntington's disease. *Neuropsychology, 13*, 525-531.

Ridenour, T. A., & Dean, R. S. (1999). Parkinson's disease and neuropsychological assessment. *International Journal of Neuroscience, 99*, 1-18.

Rochester, L., Hetherington, V., Jones, D., Nieuwboer, A., Willems, A., Kwakkel, G., & Van Wegen, E. (2004). Attending to the task: Interference effects of functional tasks on walking in Parkinson's disease and the roles of cognition, depression, fatigue, and balance. *Archives of Physical Medicine and Rehabilitation, 85,* 1578-1585.

Rogers, R. D., Sahakian, B. J., Hodges, J. R., Polkey, C. E., Kennard, C., & Robbins, T. W. (1998). Dissociating executive mechanisms of task control following frontal lobe damage and Parkinson's disease. *Brain, 121*, 815-842.

Rosen, V. M., Sunderland, T., Levy, J., Harwell, A., McGee, L., Hammond, C., Bhupali, D., Putnam, K., Bergeson, J., & Lefkowitz, C. (2005). Apolipoprotein E and category fluency: evidence for reduced semantic access in healthy normal controls at risk for developing Alzheimer's disease. *Neuropsychologia, 43*, 647-658.

Royall, D. R., Chiodo, L. K., & Polk, M. J. (2000). Correlates of disability among elderly retirees with "subclinical" cognitive impairment. *Journal of Gerontology, 55A*, 541-546.

Royall, D. R., Palmer, R., Chiodo, L., & Polk, M. (2004). Declining executive control in normal aging predicts change in functional status: The Freedom House study. *Journal of the American Geriatric Society, 52,* 346-352.

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General, 132*, 566-594.

Salthouse, T.A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology, 19*, 532-545.

Sheikh JI, Yesavage JA. (1986). Geriatric depression scale (GDS): Recent evidence and development of a shorter version. In: Brink TL, ed. *Clinical Gerontology: A Guide to Assessment and. Intervention*. Binghamton, NY: Haworth Press, pp 165-173.

Signorini, M., & Volpato, C. (2006). Action fluency in Parkinson's disease: a follow-up study. *Movement Disorders, 21*, 467-472.

Snijders, T. A. B., & Bosker, R. (1999). *Multilevel analysis*. Thousand Oaks, CA: Sage.

Storandt, M., E. A. Grant, et al. (2006). Longitudinal course and neuropathologic outcomes in original vs revised MCI and in pre-MCI. *Neurology, 67*(3): 467-473

Swanberg, M. M., Tractenberg, R. E., Mohs, R., Thal, L. J., & Cummings, J. L. (2004). Executive dysfunction in Alzheimer disease. *Archives of Neurology, 61,* 556-560.

Taylor, A.E., & Saint-Cyr, J.A. (1995). The neuropsychology of Parkinson's disease. *Brain and Cognition, 28*, 281-296.

Testa, J. A., Troster, A. I., Fields, J. A., Gleason, A. C., Salmon, D. P., & Beatty, W. W. (1998). Semantic fluency performance of patients with cortical and subcortical neurodegenerative diseases. *Aging, Neuropsychology, and Cognition, 5*, 203-214.

Troster, A. I., Fields, J. A., Testa, J. A., Paul, R. H., et al. (1998). Cortical and subcortical influences on clustering and switching in the performance of verbal fluency tasks. *Neuropsychologia, 36*, 295-304.

Troster, A. I., & Woods., S. P.(2003). Neuropsychological aspects of Parkinson's disease and Parkinsonian syndromes. In Pahwa, R., Lyons, K. E., & Koller, W.C. (Eds*.), Handbook of Parkinson's disease.* New York: Marcel Dekker.

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. (1998). Clustering and switching on verbal fluency: the effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia, 36*, 499-504.

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An update and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language, 50*, 289-335.

Author note.

This work was conducted with grant support from the Kansas City Life Sciences Institute, which is gratefully acknowledged. Additional support was provided by the Digital Electronics Core of the Center for Biobehavioral Neurosciences in Communication Disorders, grant number P30 DC-005803, for assistance with the development of the *digital ink* assessments. We especially thank Doug Kieweg for his help with this project. Portions of this research were submitted to the University of Kansas by Ellen Rozek in partial fulfillment of the requirements for the Masters degree in Psychology. Preliminary results were presented at the 2008 Cognitive Aging Conference in Atlanta. We acknowledge the assistance of Laura Berman, Carola Dopp, and Whitney McKedy in data collection and analysis.
Correspondence regarding this manuscript can be directed to Joan McDowd, jmcdowd@kumc.edu.

McDowd, J., Hoffman, L., Rozek, E., Lyons, K., Pahwa, R., Burns, J. , & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychologia*, 25, 210-225. DOI: 10.1037/a0021531. Publisher's official version: http://dx.doi.org/10.1037/a0021531. Open Access version: http://kuscholarworks.ku.edu/dspace/.

Table 1.

Means and standard deviations for cognitive tests for young adults, healthy older adults, older adults with Alzheimer's disease, older adults with Parkinson's disease.

	Young Adults		Healthy Older Adults		Older Adults with AD		Older Adults with PD		*F*
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	(3, 114)
Education	14.8_a	1.9	17.0	2.6	15.3_a	2.9	14.9_a	2.7	5.15*
MMSE	29.4_a	1.2	$28.5_{a,b}$	1.2	25.2_c	4.98	27.9_b	2.1	12.72**
GDS	1.2_a	1.2	1.1_a	1.0	1.5_a	1.5	2.8	2.9	5.48**
Boston Naming	56.0_a	3.7	55.8_a	5.2	38.1_a	13.5	54.2	4.6	35.88**
Digit Symbol	34.8_a	5.3	24.2_b	4.8	15.0_c	5.2	20.0_d	6.8	66.86**
Stroop XXXs	91.5_a	14.2	69.5_b	14.1	43.9_c	13.4	64.2_b	16.5	49.75**
Stroop Words	65.8_a	12.2	40.2_b	8.9	19.2_c	10.0	35.1_b	12.0	89.50**
Stroop Interference	0.3_a	0.1	0.4_b	0.1	0.6_c	0.2	0.4_b	0.1	20.79**
Trail A Time	46.6_a	14.1	77.7_b	30.0	123.6_c	44.1	96.0_d	31.2	31.34**
Trails B Time	55.3_a	16.4	107.4_b	59.1	199.9_c	98.1	158.5_d	87.8	21.25**
Trails Interference	-0.2_a	0.4	$-0.4_{a,b}$	0.7	-0.7_b	0.5	-0.6_b	0.5	4.13*
Forward Digits	9.8_a	2.3	$9.0_{a,b}$	3.0	7.1_c	4.6	$7.5_{b,c}$	2.4	5.11*
Backward Digits	8.2_a	3.0	6.9_b	2.5	5.3_c	1.7	$6.0_{b,c}$	1.2	9.34**
Operation Span	3.3_a	1.1	2.2_b	0.8	0.5_d	.8	1.6_c	1.1	42.12**
Letter Comparison Time	84.9_a	14.2	122.5_b	38.3	195.3_d	63.0	165.2_c	58.4	31.9**
Letter Comparison Rate	2.9_a	0.5	4.2_b	1.3	7.7_d	3.6	6.2_c	3.0	23.7**
Letter Comparison Correct	29.7_a	0.8	$28.9_{a,b}$	1.1	25.8_c	6.7	27.7_b	3.6	6.20*
Wisconsin Card Sorting Test:									
Correct Responses	52.9_a	4.4	44.4_b	10.1	30.0_d	9.4	38.2_c	11.2	33.28**
Perseverative Errors	5.5_a	1.4	8.7_b	4.8	14.4_c	6.0	13.7_c	9.4	15.10**
Nonperseverative Errors	5.6_a	3.8	10.9_b	6.1	16.6_c	7.2	11.9_b	6.7	16.50**

* $p < .05$; ** $p < .001$

Note. Results of analyses testing each subject group against each other group are represented with letters of the alphabet. Row entries sharing the same subscripts do not differ at $p < 0.05$. Row entries with different subscripts indicate significant group differences.

Table 2.

Summary (F values) from separate Group x Format ANOVAs conducted for each dependent measure used to assess fluency task performance: Correct Responses, Perseverations, Intrusions, Number of Clusters, Average Cluster Size, and Number of Cluster Switches. The Group factor compared all 4 participant groups, the Format factor compared the Traditional and Alternating Test Formats, and the Interaction tested the relation between Group and Format effects.

		Group	Format	Interaction
	degrees of freedom:	(3, 114)	(1,113)	(3, 113)
Correct Responses				
	Letters	26.18**	51.32**	2.62
	Easy Categories	33.54**	122.80**	1.75
	Hard Categories	22.29**	65.24**	0.13
	Actions	21.65**	130.60**	2.23
Perseverations				
	Letters	327.67**	2.80	2.56
	Easy Categories	431.83**	0.59	0.11
	Hard Categories	6.69*	1.89	1.11
	Actions	15.36**	2.31	0.26
Intrusions				
	Letters	0.88	10.91**	0.78
	Easy Categories	8.72*	7.31**	2.34
	Hard Categories	1.92	9.13**	0.69
	Actions	3.25*	14.51**	1.83
Number of clusters				
	Letters	6.65**	75.61**	2.39
	Easy Categories	15.10**	181.80**	0.35
	Hard Categories	16.51**	63.41**	0.05
	Actions	11.96	142.40**	2.31
Average Cluster Size				
	Letters	7.38**	399.10**	6.24**
	Easy Categories	0.61	152.31**	0.91
	Hard Categories	6.85**	255.80**	5.83**
	Actions	5.42*	118.41**	1.80
Number of Switches				
	Letters	12.52**	4.90*	0.36
	Easy Categories	19.02**	280.09**	14.00**
	Hard Categories	6.66**	12.33**	0.904
	Actions	12.02**	28.60**	1.69

* $p < .05$; ** $p < .001$

Table 3.

Results of the Stepwise Regression Analyses for Traditional Verbal Fluency for Older Adults. Only significant results are reported based on the probability of F-to enter ≤ .05.

		Step 1 (df = 1,70)			Step 2 (df = 1,69)		
		Component	R^2	*F*-to-enter	Component	R^2	*F*-to-enter
Letters							
	Correct Responses	SPD	.266	25.304	INHB	.325	6.065
	Clusters	SPD	.126	10.105			
Easy Categories							
	Correct Responses	SPD	.381	43.003	INHB	.435	6.612
	Perseverations	VRL	.142	11.556			
	Clusters	SPD	.196	17.112			
Hard Categories							
	Correct Responses	SPD	.226	20.428			
	Perseverations	VRL	.101	7.892			
	Clusters	WM	.113	8.779			
Actions							
	Correct Responses	SPD	.279	27.083	INHB	.325	4.733
	Perseverations	VRL	.123	9.808			
	Intrusions	WM	.055	4.046			
	Clusters	SPD	.139	11.280			

Note: VRL = Verbal Ability, WM = Working Memory, INHB = Inhibition, SPD = Speed

McDowd, J., Hoffman, L., Rozek, E., Lyons, K., Pahwa, R., Burns, J. , & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychologia*, 25, 210-225. DOI: 10.1037/a0021531.
Publisher's official version: http://dx.doi.org/10.1037/a0021531. Open Access version: http://kuscholarworks.ku.edu/dspace/.

Table 4.

Results of the Stepwise Regression Analyses for Alternating Verbal Fluency for Older Adults. Only significant results are reported based on the probability of F-to enter ≤ .05.

		Step 1 (df = 1,70)			Step 2 (df = 1,69)		
		Component	R^2	*F*-to-enter	Component	R^2	*F*-to-enter
Letters							
	Correct Responses	SPD	.429	52.694	INHB	.483	7.131
	Clusters	SPD	.203	17.822			
Easy Categories							
	Correct Responses	SPD	.461	59.859	INHB	.535	10.915
	Perseverations	VRL	.090	6.903			
	Clusters	SPD	.279	27.053	INHB	.360	8.765
Hard Categories							
	Correct Responses	INHB	.273	26.313	SPD	.397	14.146
	Perseverations	SPD	.128	10.274			
	Intrusions	VRL	.073	5.553			
	Clusters	INHB	.271	26.041	SPD	.351	8.504
Actions							
	Correct Responses	SPD	.298	29.694			
	Perseverations	VRL	.100	7.748			
	Clusters	WM	.215	19.210			

Note: VRL = Verbal Ability, WM = Working Memory, INHB = Inhibition, SPD = Speed

McDowd, J., Hoffman, L., Rozek, E., Lyons, K., Pahwa, R., Burns, J. , & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychologia*, 25, 210-225. DOI: 10.1037/a0021531. Publisher's official version: http://dx.doi.org/10.1037/a0021531. Open Access version: http://kuscholarworks.ku.edu/dspace/.

Table 5

Estimated Differences between Traditional versus Alternating Formats for the Intercepts and Slopes.

	Intercept Difference	Standard Error	*p*	Slope Difference	Standard Error	*P*
Letters						
Young Adults	0.275	0.232	0.242	0.169	0.070	0.021
Healthy Older Adults	1.238	0.351	0.001	-0.077	0.119	0.518
Older Adults with PD	0.057	0.351	0.871	0.151	0.119	0.207
Older Adults with AD	0.420	0.396	0.293	-0.179	0.135	0.189
Easy Categories						
Young Adults	0.685	0.281	0.020	0.235	0.093	0.017
Healthy Older Adults	0.984	0.463	0.037	0.015	0.140	0.913
Older Adults with PD	1.084	0.463	0.022	0.179	0.140	0.205
Older Adults with AD	2.349	0.521	<.0001	-0.124	0.159	0.436
Hard Categories						
Young Adults	0.225	0.645	0.729	0.388	0.194	0.055
Healthy Older Adults	0.784	0.695	0.263	0.482	0.258	0.067
Older Adults with PD	0.832	0.703	0.241	0.615	0.263	0.022
Older Adults with AD	2.353	0.805	0.005	0.368	0.314	0.245
Actions						
Young Adults	1.328	0.256	<.0001	0.318	0.092	0.002
Healthy Older Adults	2.508	0.522	<.0001	0.329	0.210	0.122
Older Adults with PD	1.615	0.527	0.003	0.717	0.213	0.001
Older Adults with AD	2.641	0.598	<.0001	0.459	0.247	0.068

Table 6

Estimated Differences between the Easy versus Hard Category Fluency Tests for the Intercepts and Slopes.

	Intercept Difference	Standard Error	*p*	Slope Difference	Standard Error	*p*
Young Adults	0.262	0.468	0.576	0.954	0.142	<.0001
Healthy Older Adults	0.607	0.515	0.240	0.734	0.157	<.0001
Older Adults with PD	1.370	0.519	0.009	0.872	0.159	<.0001
Older Adults with AD	2.431	0.587	<.0001	0.355	0.184	0.057

McDowd, J., Hoffman, L., Rozek, E., Lyons, K., Pahwa, R., Burns, J. , & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychologia*, 25, 210-225. DOI: 10.1037/a0021531. Publisher's official version: http://dx.doi.org/10.1037/a0021531. Open Access version: http://kuscholarworks.ku.edu/dspace/.

Table 7

Between-Person Effects of Verbal Ability, Processing Speed, Inhibition, Working Memory, and all 4 Predictors on Intercepts and Slopes for each Verbal Fluency Test for the 3 Groups of Older Adults (Est = model estimate; SE = standard error).

	Verbal Ability				Processing Speed				Inhibition				Working Memory			
	Traditional		Alternating		Traditional		Alternating		Traditional		Alternating		Traditional		Alternating	
	Est.	SE	Est.	SE	Est.	SE	Est.	SE	Est.	SE	Est.	SE	Est.	SE	Est.	SE
Letters																
Intercept	-.33	.19	-.20	.17	.52*	.24	.61**	.22	-.97**	.20	-.62**	.20	-1.09**	.23	-.50*	.23
Slope	-.07	.06	-.15	.07	.21*	.07	.25**	.10	-.14*	.07	-.21*	.09	-.19**	.08	-.30**	.10
Easy Categories																
Intercept	.04	.12	-.47*	.21	.11	.16	1.13**	.24	-.19	.15	-.88**	.24	-.29	.17	-.64*	.29
Slope	-.07	.05	-.12*	.06	.17**	.06	.17*	.07	-.18**	.06	-.18**	.07	-.10	.07	-.19*	.08
Hard Categories																
Intercept	-.21	.30	-.21	.35	1.07**	.39	1.03*	.43	-.84*	.36	-1.26**	.40	-1.05*	.41	-.69	.47
Slope	-.06	.09	-.06	.11	.07	.12	.28	.14	-.12	.11	-.22	.07	-.04	.13	-.19	.15
Actions																
Intercept	-.34	.19	-.71*	.32	.88**	.24	1.59**	.39	-.22	.22	-1.28**	.37	-.49	.26	-.95*	.44
Slope	-.11	.08	.13	.13	.18	.10	.00	.17	-.22*	.09	-.16	.16	-.19	.11	-.09	.18

*$p < .05$, ** $p < .01$

McDowd, J., Hoffman, L., Rozek, E., Lyons, K., Pahwa, R., Burns, J. , & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychologia*, 25, 210-225. DOI: 10.1037/a0021531. Publisher's official version: http://dx.doi.org/10.1037/a0021531. Open Access version: http://kuscholarworks.ku.edu/dspace/.

Table 8

Change in ML Deviance (-2LL) for the Conditional Models compared to Unconditional Models for the 3 Groups of Older Adults.

	Verbal Ability (df = 2)	Processing Speed (df = 2)	Inhibition (df = 2)	Working Memory (df = 2)	All 4 Predictors (df = 8)
Traditional Format					
Letters	3.69	10.75**	21.17**	21.57**	32.48**
Easy Categories	2.12	8.05*	12.13**	5.74	18.84*
Hard Categories	1.44	11.09**	9.82**	8.31*	17.34*
Actions	7.02*	22.74**	8.07*	9.20*	26.23**
Alternating Format					
Letters	6.21*	18.78**	18.73**	16.40**	30.96**
Easy Categories	10.13**	27.49**	22.46**	12.04**	42.44**
Hard Categories	0.92	13.44**	17.16**	5.16	28.27**
Actions	4.82	17.85**	17.37**	6.34*	31.41**

p * < .05; ** *p* < .01

McDowd, J., Hoffman, L., Rozek, E., Lyons, K., Pahwa, R., Burns, J. , & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychologia*, 25, 210-225. DOI: 10.1037/a0021531. Publisher's official version: http://dx.doi.org/10.1037/a0021531. Open Access version: http://kuscholarworks.ku.edu/dspace/.

Table 9

Pseudo-R^2s or Reduction in Variance after Inclusion of Between-Person Predictors for each Verbal Fluency Task for the 3 Groups of Older Adults

	Traditional Fluency					Alternating Fluency				
	R^2 Verbal Ability	R^2 Processing Speed	R^2 Inhibition	R^2 Working Memory	R^2 All 4 Predictors	R^2 Verbal Ability	R^2 Processing Speed	R^2 Inhibition	R^2 Working Memory	R^2 All 4 Predictors
Letters	19	26	26	27	31	19	27	22	24	29
Easy Categories	14	19	21	17	24	17	25	24	20	31
Hard Categories	19	21	21	20	23	31	35	34	32	38
Actions	19	24	21	20	26	20	22	23	21	25

Figure Captions.

Figure 1. Inter-word Response Times for Young Adults and 3 Groups of Older Adults on the Traditional and Alternating Letter Fluency Tasks derived from Model 4 with Heterogeneous Variance. All slopes are significantly different than zero at the .01 level or beyond.

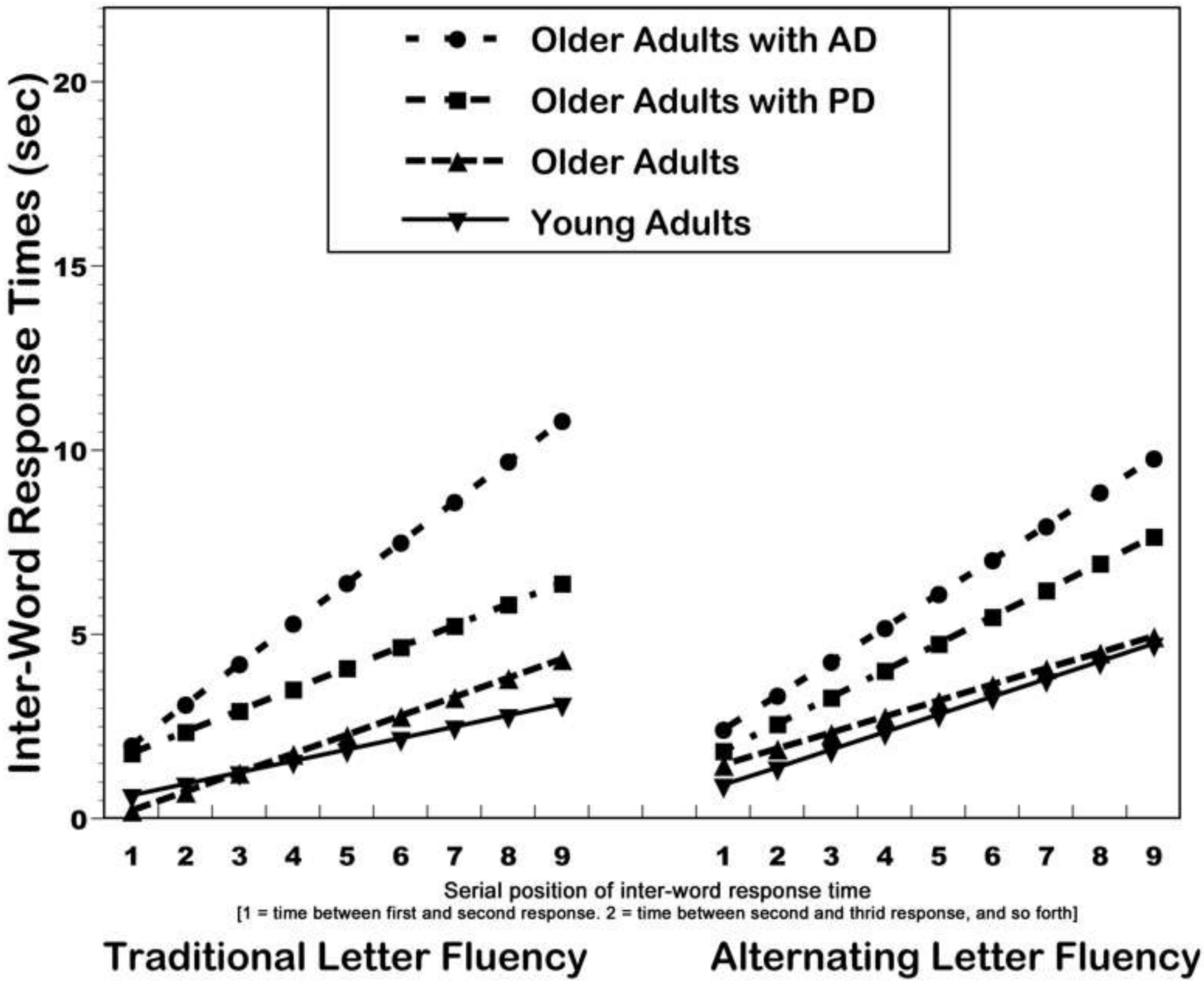
Figure 2. Inter-word Response Times for Young Adults and 3 Groups of Older Adults on the Traditional and Alternating Easy Semantic Categories Fluency Tasks derived from Model 4 with Heterogeneous Variance. All slopes are significantly different than zero at the .01 level or beyond.

Figure 3. Inter-word Response Times for Young Adults and 3 Groups of Older Adults on the Traditional and Alternating Hard Semantic Categories Fluency Tasks derived from Model 4 with Heterogeneous Variance. All slopes are significantly different than zero at the .01 level or beyond.

Figure 4. Inter-word Response Times for Young Adults and 3 Groups of Older Adults on the Traditional and Alternating Action Fluency Tasks derived from Model 4 with Heterogeneous Variance. All slopes are significantly different than zero at the .01 level or beyond.

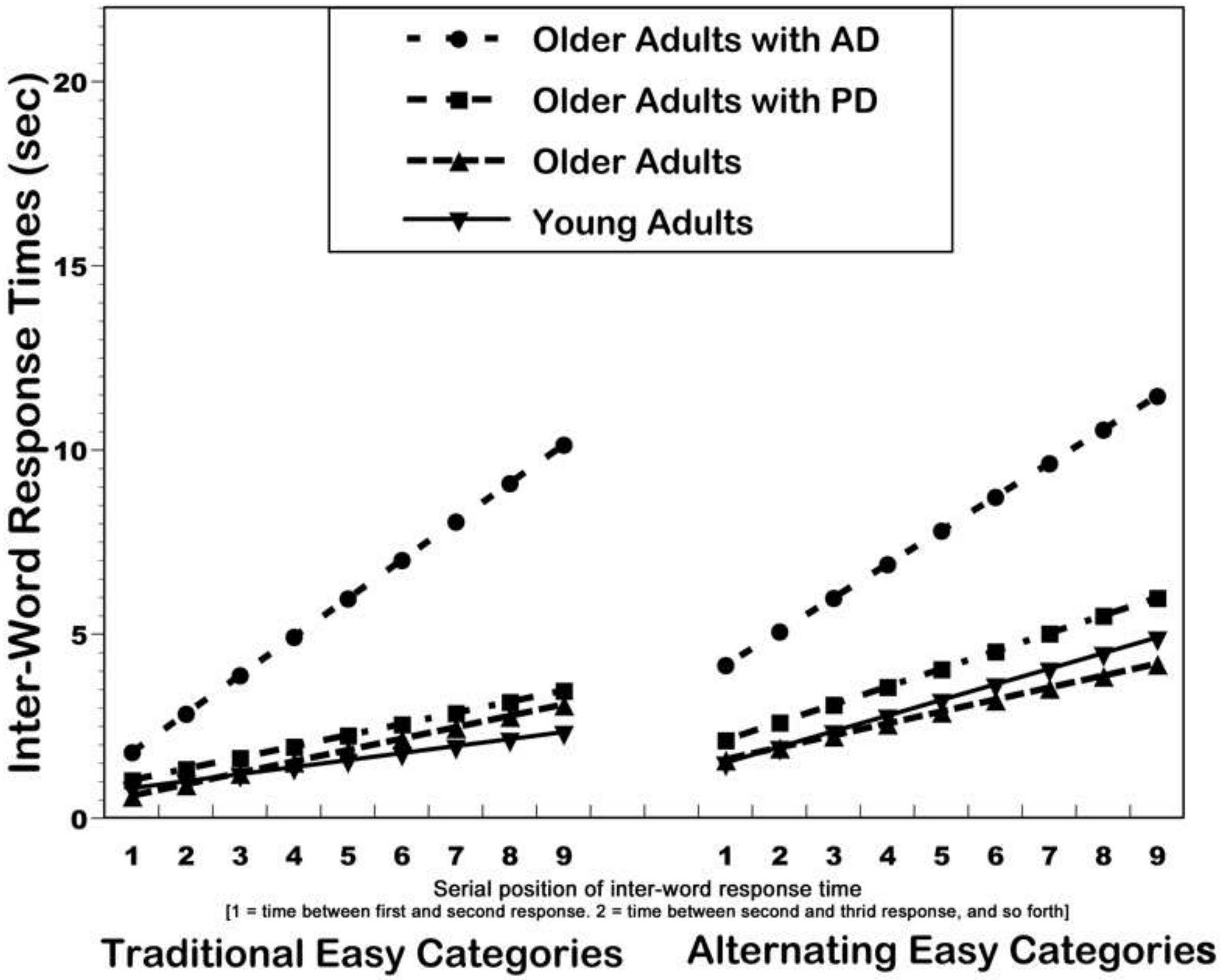
McDowd, J., Hoffman, L., Rozek, E., Lyons, K., Pahwa, R., Burns, J. , & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychologia*, 25, 210-225. DOI: 10.1037/a0021531.
Publisher's official version: http://dx.doi.org/10.1037/a0021531. Open Access version: http://kuscholarworks.ku.edu/dspace/.

FIGURE 1



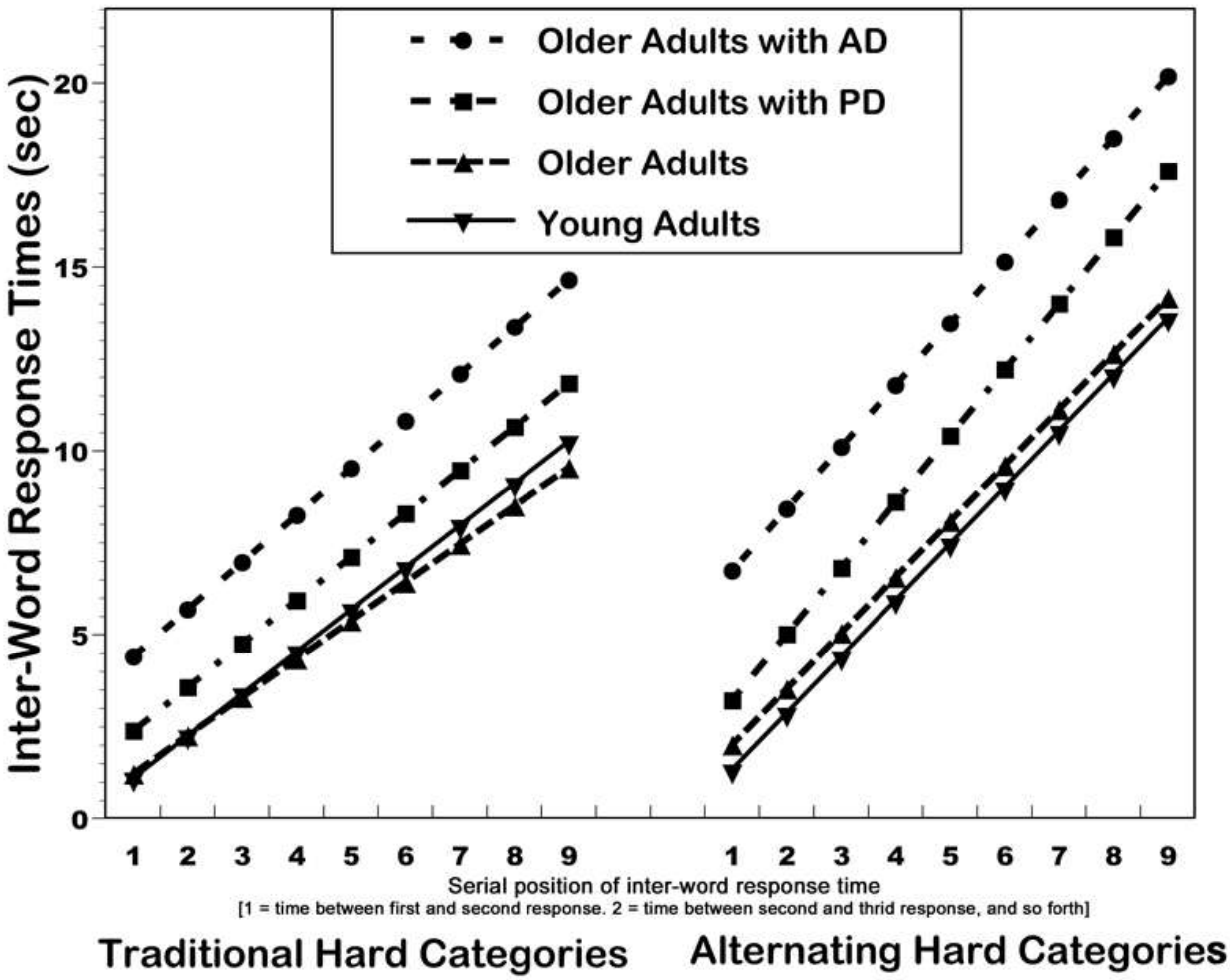
McDowd, J., Hoffman, L., Rozek, E., Lyons, K., Pahwa, R., Burns, J. , & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychologia*, 25, 210-225. DOI: 10.1037/a0021531.
Publisher's official version: http://dx.doi.org/10.1037/a0021531. Open Access version: http://kuscholarworks.ku.edu/dspace/.

FIGURE 2



McDowd, J., Hoffman, L., Rozek, E., Lyons, K., Pahwa, R., Burns, J. , & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychologia*, 25, 210-225. DOI: 10.1037/a0021531.
Publisher's official version: http://dx.doi.org/10.1037/a0021531. Open Access version: http://kuscholarworks.ku.edu/dspace/.

FIGURE 3.



McDowd, J., Hoffman, L., Rozek, E., Lyons, K., Pahwa, R., Burns, J. , & Kemper, S. (2011). Understanding verbal fluency in healthy aging, Alzheimer's disease, and Parkinson's disease. *Neuropsychologia*, 25, 210-225. DOI: 10.1037/a0021531.
Publisher's official version: http://dx.doi.org/10.1037/a0021531. Open Access version: http://kuscholarworks.ku.edu/dspace/.

FIGURE 4