Verbal fluency performance in amnestic MCI and older adults with cognitive complaints

Katherine E. Nutter-Upham, Andrew J. Saykin, Laura A. Rabin, Robert M. Roth, Heather A. Wishart, Nadia Pare, Laura A. Flashman

Abstract

Verbal fluency tests are employed regularly during neuropsychological assessments of older adults, and deficits are a common finding in patients with Alzheimer’s disease (AD). Little extant research, however, has investigated verbal fluency ability and subtypes in preclinical stages of neurodegenerative disease. We examined verbal fluency performance in 107 older adults with amnestic mild cognitive impairment (MCI, n = 37), cognitive complaints (CC, n = 37) despite intact neuropsychological functioning, and demographically matched healthy controls (HC, n = 33). Participants completed fluency tasks with letter, semantic category, and semantic switching constraints. Both phonemic and semantic fluency were statistically (but not clinically) reduced in amnestic MCI relative to cognitively intact older adults, indicating subtle changes in the quality of the semantic store and retrieval slowing. Investigation of the underlying constructs of verbal fluency yielded two factors: Switching (including switching and shifting tasks) and Production (including letter, category, and action naming tasks), and both factors discriminated MCI from HC albeit to different degrees. Correlational findings further suggested that all fluency tasks involved executive control to some degree, while those with an added executive component (i.e., switching and shifting) were less dependent on semantic knowledge. Overall, our findings highlight the importance of including multiple verbal fluency tests in assessment batteries targeting preclinical dementia populations and suggest that individual fluency tasks may tap specific cognitive processes.

Keywords: Mild cognitive impairment; Verbal fluency; Assessment; Cognition

1. Introduction

Amnestic mild cognitive impairment (MCI) is characterized by memory complaints and deficits in the absence of dementia, global cognitive decline, or significant problems with activities of daily living (ADLs) (Petersen et al., 1999). Numerous epidemiological studies have documented an accelerated rate of progression to Alzheimer’s disease (AD) in MCI patients (Tierney, Yao, Kiss, & McDowell, 2005), with annual conversion rates from about 10 to 15% per year, contrasted with the rate of 1–2% reported for healthy elderly. Over a six-year period, approximately 80% of those diagnosed with MCI will convert to AD (Petersen et al., 1999; Petersen, Doody, et al., 2001). Because the symptomatic
patients tend to show significant deficits in verbal fluency (Butters et al., 1987; Hart, 1988; Hodges, Salmon, & Butters, control” regions (Bryan, Luszcz, & Crawford, 1997; Delis & Kaplan, 2001). In comparison to healthy elderly, AD name words that begin with a specific letter, is believed to rely less on semantic knowledge stores and more on “frontal
semantic knowledge (Butters, Granholm, Salmon, Grant, & Wolf, 1987). Letter fluency, which requires the ability to rapidly name words that begin with a specific letter, is believed to rely less on semantic knowledge stores and more on “frontal control” regions (Bryan, Luszcz, & Crawford, 1997; Delis & Kaplan, 2001). In comparison to healthy elderly, AD patients tend to show significant deficits in verbal fluency (Butters et al., 1987; Hart, 1988; Hodges, Salmon, & Butters, 1991; Rosen, 1980; for review see Nebes, 1989), with greater impairment on semantic relative to letter fluency tasks (Chan, Butters, Salmon, & McGuire, 1993; Martin & Fedio, 1983). Investigation of the discriminative value of fluency subtypes in AD has revealed that semantic fluency (100% sensitivity, 92.5% specificity) is superior to letter fluency (89% sensitivity, 85% specificity) in predicting group membership (Monsch et al., 1992).

Several explanations have been proposed to account for the discrepancy between performance on tests of semantic and letter fluency in AD. One model posits that AD patients suffer from impaired retrieval from an otherwise intact semantic memory store (Nebes & Brady, 1988; Nebes, Martin, & Horn, 1984). In this conceptualization, associations between attributes of a given semantic store (e.g., elephant attributes might include, heavy, big, wild) are intact, and deficits are entirely due to slowed retrieval. This is considered to be a non-linguistic or executive control deficit (Henry, Crawford, & Phillips, 2004). A second model attributes impaired semantic verbal fluency to a loss (Hodges, Salmon, & Butters, 1992; Monsch et al., 1992; Salmon, Butters, & Chan, 1999) or disorganization (Grober, Buschke, Kawas, & Fuld, 1985) in the actual structure or content of semantic knowledge. Loss of stored representations or attributes is thought to slow word production.

Empirical investigations of the proposed models have been conducted by comparing patients’ performance on various verbal fluency tasks. When AD patients were asked to rapidly name supermarket items on the Mattis Dementia Rating Scale they produced many super-ordinate responses but few exemplars (Martin & Fedio, 1983). For example, participants were able to name categories (i.e., meat) but unable to give specific exemplars of that category (i.e., lamb, beef). Inability to access exemplars was considered to be evidence for the disruption of semantic memory stores. Participants in this study could access the memory store but all the information believed to be contained in the store was not present. Henry and Crawford (2004) examined focal lesion cohorts and found that severity of frontal lobe injuries was associated with both letter (r = .52) and semantic (r = .54) fluency, while severity of temporal lobe lesions showed a differential association with letter (r = .44) and semantic (r = .61) fluency. These findings point to the use of frontal regions in both tasks, but highlight the contribution of temporal lobe regions for semantic fluency. Research also has shown that AD patients experience a loss of semantic knowledge more generally (Norton, Bondi, Salmon, & Goodglass, 1997), and they tend to produce more semantic-related errors on confrontation naming tasks as compared to healthy elders (Bayles & Tomoeda, 1983).

AD also has been shown to be associated with a distortion in the organization of semantic memory. For example, using cognitive maps Chan et al. (1993) observed that AD patients used attributes less consistently and categorized more often on concrete criteria as compared to normal healthy elderly. Furthermore, functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) studies have revealed anomalous activity in patients with early AD during semantic and episodic memory processing including spatially extended areas of activation (Becker, Lopez, & Butters, 1996; Becker, Mintun, et al., 1996; Saykin et al., 1999).
participants with AD allocate more brain tissue to complete a semantic task as compared with older adult controls. Although much of the available research appears to support the idea that increased difficulty in semantic relative to letter fluency is related to a genuine loss of semantic knowledge rather than a deficit in retrieval from an intact semantic store, it is unclear if this is the case in early preclinical stages.

Researchers have yet to systematically investigate performance on verbal fluency subtypes in older adults with MCI. Standish, Molloy, Cunje, and Lewis (2007) found that a subtest of verbal fluency (type not specified, within a 135-item cognitive screening test) was best able to discriminate between healthy older adults and those with MCI after controlling for age and education. Saxton et al. (2004) included verbal fluency tasks in research on neuropsychological performance approximately 1.5–5 years prior to the onset of AD. Results indicated that decline in category (semantic) fluency was one of the predictors of subsequent conversion to AD in older adults. Another study investigated practice effects on category fluency over a one-week interval, and found that amnestic MCI and AD participants did not benefit from repeated exposure to the task while healthy elders did benefit (Cooper, Lacritz, Weiner, Rosenberg, & Cullum, 2004) suggesting a decline in category fluency performance for these populations. Recent work by Murphy, Rich, and Troyer (2006) found that amnestic MCI participants produced significantly fewer words on category (semantic), but not phonemic fluency when compared to healthy controls, though overall performance on both tasks fell within normal limits clinically. The researchers concluded that MCI patients experienced difficulties making semantic associations between exemplars of subcategories rather than difficulties searching through lexical representations, suggesting degeneration of brain areas beyond that of the hippocampus (i.e., inclusion of cortical areas supporting semantic memory) at this early disease stage. Overall, extant research suggests a possible decline in verbal fluency for individuals with amnestic MCI, particularly apparent on semantic fluency subtests.

The goal of the current study was to systematically investigate performance on various subtypes of verbal fluency in older adults with amnestic MCI, older adults with significant cognitive complaints (CC) despite neuropsychological test performance within normal limits, and healthy matched controls (HC). We evaluated the differential ability of the various fluency measures to detect differences among our participant groups. The CC group is of special interest in light of recent research suggesting that these individuals show brain changes intermediate between those seen in MCI and healthy older adults without complaints, and may represent a “pre-MCI” stage (Saykin et al., 2006). Another goal was to explore the underlying constructs of fluency measures to clarify the relative contributions of executive functioning and semantic knowledge on performance. While marked fluency deficits are not typically expected in preclinical dementia, we hypothesized the presence of subtle difficulties that could relate to organization of the attributes of semantic stores or other aspects of executive function. In addition to commonly used tests of semantic and letter fluency, we included other subtypes that have yet to be investigated with regard to the contributions of semantic knowledge versus executive control features. These include switching tasks, which require participants to alternate between semantic categories (and more explicitly involve executive functions) and action/verb naming tasks, also believed to be mediated by frontal regions (Baxter & Warrington, 1985; Cappa, Sandrini, Rossini, Sosta, & Miniussi, 2002; Miozzo, Soardi, & Cappa, 1994; Piatt, Fields, Paolo, & Troster, 2004).

2. Methods

2.1. Recruitment

37 non-depressed older adults diagnosed with amnestic MCI, and 33 demographically matched healthy older adult controls (HC). We also included a third group of 37 non-depressed older adults who presented with significant cognitive complaints but who performed within normal limits on neuropsychological testing. Participants were enrolled consecutively in a longitudinal study of preclinical and early dementia, through comprehensive screening including a standardized phone interview (Rabin et al., 2007) and review of medical records. Participants were recruited from flyers, newspaper ads, public lectures, and referrals from Dartmouth Hitchcock Medical Center clinics. Enrollment criteria included: at least 60 years of age, right-handed, fluent in English and a minimum of 12 years of formal education. Participants were excluded for any medical, psychiatric, or neurological conditions (other than MCI) that could affect brain structure or cognition. History of head trauma with loss of consciousness for more than 5 min and current or past substance abuse also served as exclusionary criteria. Participants provided written informed consent according to procedures approved by the Institutional Committee for the Protection of Human Subjects.
2.2. Neuropsychological assessment and group classification

Participants underwent comprehensive neuropsychological evaluation including measures of memory, attention, executive function, processing speed, language, general intellectual ability and standard dementia screens. Tests included: verbal fluency subtests of the Delis-Kaplan Executive Function System (D-KEFS, Delis & Kaplan, 2001); Action Naming Test (Piatt, Fields, Paolo, & Troster, 1999); Dementia Rating Scale, Second Edition (DRS-2, Jurica, Leitlen, & Mattis, 2001); Mini Mental State Examination (MMSE, Folstein, Folstein, & McHugh, 1975); California Verbal Learning Test, Second Edition (CVLT-II, Delis, Kramer, Kaplan, & Ober, 2000); Boston Naming Test (BNT, Goodglass, Kaplan, & Barresi, 2001); D-KEFS Trail Making Test; Wechsler Adult Intelligence Scale, Third Edition (WAIS-III, Digit Symbol, Digit Span, Block Design, Vocabulary, and Information subtests, Wechsler, 1997); Wechsler Memory Scale, Third Edition (WMS-III, Logical Memory and Visual Reproduction subtests, Wechsler, 1997); and Wisconsin Card Sorting Test (WSCT short form; Heaton, Chelune, Talley, Kay, & Curtiss, 1993). Estimates of baseline intellectual functioning included the American National Adult Reading Test (ANART; Grober, Sliwinski, Schwartz, & Saffran, 1989) and the Barona Index (Barona, Reynolds, & Chastian, 1984). Level of cognitive complaint was assessed using self- and informant-report questionnaires including: Memory and Aging Telephone Screen (Rabin et al., 2007); Memory Self-Rating Questionnaire (Squire, Wetzel, & Slater, 1979), Neurobehavioral Function/Activities of Daily Living Scale (NBF-ADL self- and informant-versions; Saykin et al., 1991), Memory Assessment Questionnaire (Pfeffer, Kurosaki, Harrah, Chance, & Filos, 1982; Santulli et al., 2005), Informant Questionnaire on Cognitive Decline in the Elderly (self- and informant-versions; Jorm, Scott, & Jacomb, 1989), and cognitive items from the Geriatric Depression Scale (GDS; Yesavage et al., 1983). A Cognitive Complaint Index ranging from 0 to 100 was calculated based on the total number of items that could be endorsed across all subjective measures (for details see Saykin et al., 2006). A Board-certified geropsychiatrist conducted a semi-structured interview to rule out depression or other psychiatric disorders. Blood samples were obtained to determine Apolipoprotein E (ApoE) genotype, and participants underwent structural brain MRIs, which were reviewed by a Board-certified neuroradiologist to rule out incidental pathology.

All data were derived from participants’ baseline evaluation in our longitudinal study, with the exception of five participants who received earlier versions of the fluency tasks; for these individuals, we used data from the second study visit (approximately one year following baseline). The group membership of the five participants whose second study visit data were utilized was as follows: two HCs, two CCs, and one MCI.

Group classifications were made following careful review of neuropsychological assessment results, self- and informant-report data, and geropsychiatric evaluation. Group classification was determined by a consensus panel review using Petersen workgroup criteria for amnestic MCI (Petersen, Smith, et al., 2001). Participants diagnosed as MCI had: (1) preserved general cognitive functioning (as measured by cutoff scores on the DRS-2 and MMSE); (2) generally normal activities of daily living (as reported on self- and informant-versions of the NBF-ADL scale); (3) no dementia; (4) significant memory complaints (endorsed at least 20% of possible complaints across all inventories or complaints deemed significant by clinical consensus); and (5) impaired memory (memory scores 1.5 S.D. below the mean established for age- and education-matched controls). Participants classified as CC had: (1) preserved general cognitive functioning; (2) generally normal activities of daily living; (3) no dementia; (4) significant memory complaints; and (5) intact memory ability. Participants classified as HC had: (1) preserved general cognitive functioning; (2) generally normal activities of daily living; (3) no dementia; (4) no significant cognitive complaints; and (5) intact memory ability. The Peterson workgroup criteria for MCI dictate that participants have preserved general cognition and isolated memory problems; notable executive or language impairment (including verbal fluency) was grounds for exclusion.

Select demographic and neuropsychological test variables are presented in Table 1. The average age of participants was 72.3 (S.D. = 6.0) years and the average level of education was 16.7 (S.D. = 2.8) years. All other key demographic variables were shown to be insignificant. Though no participant was clinically depressed or scored in the depressed range on the adjusted GDS (excludes the four cognitive items), CC and MCI participants tended to endorse several more items than HC. As expected based on study classification criteria, scores on the DRS-2 and MMSE were reduced in MCI relative to the other groups, though MCI participants scored above the cutoff for dementia. MCI participants also showed significant deficits on tests of memory (e.g., DRS-2, CVLT-II, WMS-III) relative to HC. Based on study classification criteria, the Cognitive Complaint Index was significantly elevated in both MCI and CC relative to HC.
Table 1
Participant demographics and neuropsychological test data

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Participant group</th>
<th>p</th>
<th>Group differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HC (n = 33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M (S.D.)</td>
<td>M (S.D.)</td>
<td>M (S.D.)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>71.09 (5.51)</td>
<td>73.35 (5.98)</td>
<td>71.76 (6.66)</td>
</tr>
<tr>
<td>Education (years.)</td>
<td>17.27 (2.43)</td>
<td>16.41 (2.91)</td>
<td>16.24 (2.88)</td>
</tr>
<tr>
<td>Gender (M, F)</td>
<td>11, 22 –</td>
<td>16, 21 –</td>
<td>17, 20 –</td>
</tr>
<tr>
<td>GDS (/26)</td>
<td>1.85 (2.412)</td>
<td>4.24 (4.02)</td>
<td>3.78 (3.09)</td>
</tr>
<tr>
<td>DRS-2 (/144)</td>
<td>140.9 (2.05)</td>
<td>140.9 (2.59)</td>
<td>137 (4.24)</td>
</tr>
<tr>
<td>MMSE, total score (/30)</td>
<td>29.15 (.91)</td>
<td>29 (1.11)</td>
<td>26.59 (2.1)</td>
</tr>
</tbody>
</table>

Note: Data are means, standard deviations, and raw scores except where indicated. HC, healthy control; CC, cognitive complaint; MCI, mild cognitive impairment; GDS, Adjusted Geriatric Depression Scale (n/26 non-cognitive items); DRS-2, Dementia Rating Scale (Supermarket Items); MMSE, Mini Mental State Exam.

2.3. Fluency tasks

All participants were administered five verbal fluency tasks as part of a larger test battery, each with a 60 s time limit. Responses were given orally and recorded by hand by highly trained technicians or postdoctoral fellows. Three of the fluency tasks were D-KEFS subtests, and required participants to rapidly generate words that: (1) begin with a certain letter (letter fluency), (2) belong to a specific category (semantic fluency), or (3) alternate between two different semantic categories (switching fluency). Participants also completed (4) action naming fluency, which required them to name things “that people do” (i.e., verbs). The final task: (5) supermarket naming fluency, was a subtest of the DRS-2, and required participants to name items they could “find or buy in a supermarket.” All fluency scores were derived by summing correct utterances and excluded repetitions. For the DRS-2 supermarket naming test, the maximum number of points awarded was 20 (as per the standardized instruction manual). For the D-KEFS subtests, participants were randomly administered either the standard or alternate test form to reduce practice effects engendered by multiple administrations of the same items. Alternate test forms were not available for the action naming or supermarket item fluency tasks. Form equivalency has been demonstrated by previous research (Cunje, Molloy, Standish, & Lewis, 2007; Delis & Kaplan, 2001), and our analysis yielded insignificant results: \( \chi^2(2, N = 107) = .708, p > .05 \). Therefore the different versions of the fluency scores were collapsed into four variables for all participants: “letter,” “category,” “switching,” and “shifting.” Letter fluency scores were a combination of the three individual letter trials (FAS or BHR), category scores were a composite of the two individual trials (animals and boys or clothing and girls), and switching scores comprised a single trial with two categories (fruits and furniture or instruments and vegetables). Number of correct shifts between semantic categories also was analyzed. In all, six fluency variables were included: letter, category, switching, shifting, action naming, and supermarket naming.

2.4. Statistical procedures

Skew and kurtosis statistics were calculated for each fluency measure to explore the shape of the distributions. Supermarket naming was highly negatively skewed (statistic = −3.19, standard error = .23), while all other fluency measures where normally distributed. A Kruskal–Wallis was conducted on the supermarket items using the conservative .01 cutoff as a cautionary approach to non-parametric testing. One-way multivariate analysis of variance (MANOVA) was conducted to evaluate whether the groups (MCI, CC, HC) showed an overall difference on the fluency measures, as well as to provide control over Type I error. MANOVA was followed by analyses of variance (ANOVA)s and Tukey’s HSD post-hoc tests to identify the loci of group differences for each measure. Significance level was set at \( p < .05 \) for the MANOVA and for the ANOVA and post-hoc tests. Eta-squared was employed as the measure of effect size. In order to evaluate the underlying constructs of multiple verbal fluency measures, a factor analysis with varimax rotation was used. Number of factors extracted was determined using a minimum eigenvalue of one criterion. In order to investigate the classification value of our factors, we conducted discriminant analyses using the MCI and HC participants. We also examined correlations among fluency scores and their relation with select neuropsychological variables (using Pearson correlation coefficients).
Table 2
Verbal fluency scores by diagnostic group

<table>
<thead>
<tr>
<th>Fluency task</th>
<th>Participant group</th>
<th>p</th>
<th>η²</th>
<th>Group differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HC (n = 33)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>S.D.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letter</td>
<td>44.52</td>
<td>10.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>42.12</td>
<td>6.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switching</td>
<td>14.85</td>
<td>2.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shifting</td>
<td>13.82</td>
<td>2.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DRS-2</td>
<td>19.94</td>
<td>.348</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Action naming</td>
<td>20.97</td>
<td>4.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC (n = 37)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>S.D.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letter</td>
<td>44.14</td>
<td>10.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>41.03</td>
<td>9.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switching</td>
<td>14.49</td>
<td>2.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shifting</td>
<td>13.76</td>
<td>3.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DRS-2</td>
<td>19.65</td>
<td>.919</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Action naming</td>
<td>19.05</td>
<td>4.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MCI (n = 37)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>S.D.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letter</td>
<td>37.49</td>
<td>10.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>35.38</td>
<td>6.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switching</td>
<td>11.54</td>
<td>2.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shifting</td>
<td>10.54</td>
<td>2.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DRS-2</td>
<td>19.30</td>
<td>1.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Action naming</td>
<td>17.84</td>
<td>4.67</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: HC, healthy control; CC, cognitive complaint; MCI, mild cognitive impairment. p values and η² values were obtained from ANOVA results; group differences were obtained from Tukey’s HSD post hoc analyses.

3. Results

3.1. Group differences

Descriptive statistics for the fluency measures are presented in Table 2. The overall multivariate test (Wilk’s Lambda) was significant \(F(2, 104) = 4.16, p < .001\), and follow-up ANOVAs indicated that the groups differed with respect to letter \(F(2, 104) = 5.20, p < .01, \eta^2 = .091, \text{power} = .82\), category \(F(2, 104) = 7.77, p < .001, \eta^2 = .130, \text{power} = .95\), switching \(F(2, 104) = 16.90, p < .001, \eta^2 = .245, \text{power} = 1.00\), shifting fluency \(F(2, 104) = 14.56, p < .001, \eta^2 = .219, \text{power} = 1.00\), and action naming \(F(2, 104) = 4.24, p = .017, \eta^2 = .075, \text{power} = .73\). We repeated the MANOVA and subsequent ANOVAs excluding the five participants, whose second study visit data were utilized, with no change in the pattern of results.

Due to the high degree of skewness in the supermarket naming, a Kruskal–Wallis test was performed, but results did not reach the .01 cutoff value: \(\chi^2(2, N = 107) = 7.521, p < .01\). Post hoc analyses to the univariate ANOVAs using the Tukey HSD correction revealed that MCI participants scored significantly lower than CC and HC on letter, category, switching, and action naming tasks. We repeated the multivariate analyses with adjusted GDS score as a covariate, with the same overall pattern of findings. Results indicated that depressive symptoms (i.e., adjusted GDS score) did not account for a significant amount of the variance between groups (3% of the variance; \(p > .05\)).

3.2. Factor analysis of fluency tasks

Factor analysis was employed to investigate the heterogeneity of fluency tasks with a cutoff eigenvalue of 1. Table 3 provides the factor loadings for the cognitive measures. Two significant factors emerged, which we termed Switching and Production. Switching accounted for 56.1% of the variance while the Production factor accounted for 22.8%. In all, 79% of the variance in fluency performance was explained by these two factors. Switching included switching and shifting fluency tasks. Production had loadings from letter, category, and action naming tasks. It is worth noting that

Table 3
Factor loadings: Switching and Production

<table>
<thead>
<tr>
<th>Fluency tasks</th>
<th>Switching</th>
<th>Production</th>
</tr>
</thead>
<tbody>
<tr>
<td>Letter</td>
<td>.174</td>
<td>.804</td>
</tr>
<tr>
<td>Category</td>
<td>.280</td>
<td>.803</td>
</tr>
<tr>
<td>Action naming</td>
<td>.995</td>
<td>.764</td>
</tr>
<tr>
<td>Switching</td>
<td>.960</td>
<td>.236</td>
</tr>
<tr>
<td>Shifting</td>
<td>.971</td>
<td>.180</td>
</tr>
</tbody>
</table>

Note: Items in bold load together on a given factor.
the effect sizes for the individual tasks were adequate for the Switching component tasks and poor for the Production component tasks (Table 2).

To investigate the ability of these two factors to discriminate groups, a group-by-factor analysis was conducted. Factor scores were calculated by taking participants’ Z-scores for each of the five fluency tasks, multiplied by the amount of variance accounted for on the given factor, and summing the product of these variables. The resulting values were Z-scores on each factor, accounting for the variance described by each individual test; both factor scores were normally distributed (Production factor skewness statistic = -.099; Switching factor skewness statistic = .056).

Notably, Z-scores were calculated using the entire sample (instead of HC exclusively), ensuring a conservative approach. Results revealed a significant difference for both the Switching \( F(2, 104) = 12.47, p < .001, \eta^2 = .197 \) and Production \( F(2, 104) = 4.56, p = .013, \eta^2 = .081 \) factors (see Fig. 1). Post hoc analysis using Tukey’s HSD correction revealed that MCI participants were significantly lower than CC and HC on the Switching factor, and that MCI participants were significantly lower than HC on the Production factor.

Two discriminant function analyses were conducted to determine which factor best predicted membership as an MCI or HC. For the Switching factor, the discriminant analysis yielded a significant Wilk’s lambda \( \Lambda = .78, \chi^2(1, N = 70) = 16.66, p < .01 \), and correctly classified 24 of 33 HCs and 25 of 37 MCIs. Overall, the Switching factor correctly classified 75% of the participants. The discriminant analysis for the Production factor also yielded a significant Wilk’s lambda \( \Lambda = .88, \chi^2(1, N = 70) = 8.36, p < .01 \), and correctly classified 20 of the 33 HCs and 22 of 37 MCIs. Overall, the Production factor correctly classified 60% of the participants.

3.3. Neuropsychological correlates

A final set of analyses examined relations among fluency tasks and between fluency tasks and select neuropsychological variables. Pearson correlation coefficients (\( p < .01 \) level) revealed statistically significant relations among all six fluency measures (range = .23–.96; Table 4), with the exception of supermarket naming and action naming. Statistically significant, moderate correlations were observed between fluency measures and various neuropsychological test scores (Table 5). Of particular interest are relations between fluency measures and tests known to represent verbal knowledge (WAIS-III Vocabulary, Information, and the BNT) or executive function (WCST, WAIS-III Digit Symbol, and D-KEFS Trail Making Test, Switching condition). WAIS-III Vocabulary, WAIS-III Information, and the BNT were correlated with action naming (\( r = .54, .45, .30 \), respectively) and letter fluency (\( r = .52, .38, .33 \), respectively). WAIS-III Vocabulary and the BNT were correlated with category fluency (\( r = .33, .27 \), respectively), though WAIS-III Information was not. Switching and shifting tasks were not significantly correlated with WAIS-III Vocabulary, WAIS-III Information, or the BNT (all \( r \) values \( \leq .24 \)). The WCST (total categories and number of perseverative errors) did not correlate with any fluency measure, but WAIS-III Digit Symbol correlated with action naming (\( r = .54 \)), letter (\( r = .40 \)), category (\( r = .40 \)), switching (\( r = .41 \)), and shifting (\( r = .40 \)) scores.
Table 4
Correlations among fluency tasks

<table>
<thead>
<tr>
<th></th>
<th>Letter</th>
<th>Category</th>
<th>Switching</th>
<th>Shifting</th>
<th>DRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Action naming</td>
<td>.484*</td>
<td>.474*</td>
<td>.302*</td>
<td>.285*</td>
<td>.156</td>
</tr>
<tr>
<td>Letter</td>
<td>–</td>
<td>.535*</td>
<td>.367*</td>
<td>.302*</td>
<td>.263*</td>
</tr>
<tr>
<td>Category</td>
<td>–</td>
<td>–</td>
<td>.424*</td>
<td>.379*</td>
<td>.246*</td>
</tr>
<tr>
<td>Switching</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>.956*</td>
<td>.233*</td>
</tr>
<tr>
<td>Shifting</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>.273*</td>
</tr>
</tbody>
</table>

*p < .01.

Table 5
Fluency task correlations with neuropsychological tasks

<table>
<thead>
<tr>
<th></th>
<th>Letter</th>
<th>Category</th>
<th>Switching</th>
<th>Shifting</th>
<th>Action naming</th>
</tr>
</thead>
<tbody>
<tr>
<td>WAIS-III Vocabulary</td>
<td>.52**</td>
<td>.33**</td>
<td>.22*</td>
<td>.19</td>
<td>.54**</td>
</tr>
<tr>
<td>WAIS-III Digit Symbol</td>
<td>.40**</td>
<td>.40**</td>
<td>.41**</td>
<td>.40**</td>
<td>.45**</td>
</tr>
<tr>
<td>WAIS-III Information</td>
<td>.38**</td>
<td>.24*</td>
<td>.17</td>
<td>.15</td>
<td>.45**</td>
</tr>
<tr>
<td>D-KEFS Trail Making</td>
<td>−.28**</td>
<td>−.30**</td>
<td>−.32**</td>
<td>−.28**</td>
<td>−.29**</td>
</tr>
<tr>
<td>Boston Naming Test</td>
<td>.33**</td>
<td>.27**</td>
<td>.25*</td>
<td>.21*</td>
<td>.30**</td>
</tr>
</tbody>
</table>

*p < .05.

**p < .01.

4. Discussion

4.1. Overview

Amnestic MCI is conceptualized as memory complaints and episodic memory impairment in the absence of dementia or global cognitive decline. Recent research has indicated that subtle decline in other areas of cognition, such as executive functioning, may coexist with episodic memory deficits. Researchers have included verbal fluency in predictive models for conversion to AD, raising the question of when verbal fluency becomes compromised in the course of disease. Results of the current study revealed an overall decline in verbal fluency performance in patients diagnosed with amnestic MCI relative to demographically matched healthy controls. Additionally, factor analytic results suggested that individual fluency tasks may tap specific cognitive processes, and correlational findings implied that all fluency tasks involved executive control to some degree, while those with an added executive component (i.e., switching and shifting) were less dependent on semantic knowledge.

4.2. Group differences

Examination of group differences revealed that MCI participants scored lower on most fluency tasks (i.e., letter, category, switching, shifting) as compared to HC. In contrast to recent findings by Murphy et al. (2006), which showed reduced semantic fluency and intact phonemic fluency, our MCI group showed reduced fluency (relative to healthy controls) on both phonemic and semantic tasks. Future research should further explore the degree to which various fluency subtypes are selectively compromised in MCI and the relation between this compromise and the underlying cognitive demands of the tasks in question. Another important study finding was that mean scores across all fluency tasks followed the pattern of HC > CC > MCI. As with previous findings from our laboratory (Saykin et al., 2006), CCs showed normal performance despite significant complaints about aspects of cognition. These individuals are known to manifest subtle structural and functional brain changes (Saykin et al., 2004, 2006), and may be a prodrome of MCI. It therefore will be important to follow this at-risk cohort longitudinally to determine when, and under what conditions, fluency deficits arise.

1 Refer to Saykin et al. (2006) for further details regarding this cohort [overlap > 86% with participants in the current study].
Table 6
D-KEFS scaled scores by participant group

<table>
<thead>
<tr>
<th>Participant group</th>
<th>M</th>
<th>S.D.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>D-KEFS letter</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HC</td>
<td>12.8485</td>
<td>3.08344</td>
<td>33</td>
</tr>
<tr>
<td>CC</td>
<td>12.8649</td>
<td>3.00150</td>
<td>37</td>
</tr>
<tr>
<td>MCI</td>
<td>11.0270</td>
<td>3.10454</td>
<td>37</td>
</tr>
<tr>
<td>Total</td>
<td>12.2243</td>
<td>3.15723</td>
<td>107</td>
</tr>
<tr>
<td><strong>D-KEFS category</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HC</td>
<td>13.4545</td>
<td>2.70521</td>
<td>33</td>
</tr>
<tr>
<td>CC</td>
<td>12.8649</td>
<td>3.33468</td>
<td>37</td>
</tr>
<tr>
<td>MCI</td>
<td>11.1081</td>
<td>2.85564</td>
<td>37</td>
</tr>
<tr>
<td>Total</td>
<td>12.4393</td>
<td>3.12379</td>
<td>107</td>
</tr>
<tr>
<td><strong>D-KEFS switching</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HC</td>
<td>13.2727</td>
<td>3.14516</td>
<td>33</td>
</tr>
<tr>
<td>CC</td>
<td>12.8649</td>
<td>3.10163</td>
<td>37</td>
</tr>
<tr>
<td>MCI</td>
<td>9.4595</td>
<td>2.85379</td>
<td>37</td>
</tr>
<tr>
<td>Total</td>
<td>11.8131</td>
<td>3.46446</td>
<td>107</td>
</tr>
<tr>
<td><strong>D-KEFS shifting</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HC</td>
<td>12.9394</td>
<td>3.09172</td>
<td>33</td>
</tr>
<tr>
<td>CC</td>
<td>12.9459</td>
<td>2.97159</td>
<td>37</td>
</tr>
<tr>
<td>MCI</td>
<td>9.7568</td>
<td>2.74272</td>
<td>37</td>
</tr>
<tr>
<td>Total</td>
<td>11.8411</td>
<td>3.27987</td>
<td>107</td>
</tr>
</tbody>
</table>

Note: HC, healthy control; CC, cognitive complaint; MCI, mild cognitive impairment. D-KEFS scaled scores are based on mean = 10 and standard deviation = 1.

Although statistically significant group differences were obtained across fluency measures, mean MCI scores generally fell within the low average range clinically (i.e., approx. 1 S.D. below the mean of HC). Verbal fluency measures thus appear to be sensitive to subtle, non-amnestic decline in MCI, and these findings could have implications for everyday functioning. For example, research has shown that verbal fluency performance makes a significant independent contribution to the prediction of instrumental activities of daily living as reported by caregivers of community-dwelling older adults (Cahn-Weiner, Boyle, & Malloy, 2002). In particular, that study showed that response generation and behavioral persistence (measured by verbal fluency) were strongly related to functional status in everyday tasks in which older adults were given little external structure. Longitudinal follow of our MCI cohort will be important to determine the predictive utility of various fluency measures for conversion to AD. It is also important to consider findings from our discriminant function analyses, which revealed the moderate diagnostic classification value of our two factors (60% and 75% overall for MCI and HC), with relatively stronger support for Switching as compared to Production. Although fluency measures will never replace tests of episodic memory as the most sensitive measures for early detection of preclinical AD, they may find incorporation in multivariate regression models meant to identify the optimal test combination for classification.

4.3. Heterogeneity of fluency tasks

To our knowledge, this was the first study to incorporate tests of switching fluency (alongside category and letter fluency) in the comprehensive neuropsychological assessment of MCI. Our factor analytic results yielded two interpretable factors: Production (letter, category, action naming) and Switching (switching, shifting), partially supporting the idea that individual fluency tasks rely on different cognitive processes. As previously noted, phonemic fluency depends on search strategies involving lexical representations while category fluency relies more on semantic memory and requires search strategies that are semantic extensions of a target superordinate, (e.g., Delis & Kaplan, 2001; Henry

2 This analysis also was performed using the D-KEFS scaled scores and the same pattern of findings emerged. Scaled score means and standard deviations for the three groups are reported in Table 6.
et al., 2004). Furthermore, lesion studies have identified the temporal lobes as the primary loci for category fluency, while the frontal lobes as more influential in letter fluency (Henry & Crawford, 2004; Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998). Action or verb naming also has been found to have a strong executive component, possibly mediated by frontal regions (Baxter & Warrington, 1985; Cappa et al., 2002; Damasio & Tranel, 1993; Miozzo et al., 1994).

We initially hypothesized that all fluency tasks known to possess a strong executive component (i.e., switching, action naming, letter fluency) would group together upon factor analysis. Instead, our results indicated that tasks with a switching or shifting component loaded together, while remaining tasks loaded on a second factor, labeled Production. A possible explanation for this finding is that the two tasks rely on different cognitive processes. While the Switching tasks provide participants with a lexical search cue, the Production tasks do not. This idea is supported by findings by Baldo, Shimamura, Delis, Kramer, and Kaplan (2001) who showed that participants with frontal lobe lesions did not exhibit impairment of verbal fluency switching—possibly related to the provision of exogenous cuing associated with this task. One might speculate that the Production factor will engage certain brain regions (e.g., inferior frontal regions involved in articulation of semantic memory), whereas the Switching component could reflect involvement of other areas (e.g., dorsal lateral prefrontal cortex). Investigation of participants’ performance of these tasks while undergoing fMRI or analysis of correlations with structural data using voxel-based morphometric techniques might elucidate the cognitive underpinnings of fluency tasks with a switching component. It also will be important to utilize data-reduction techniques within a larger group of healthy controls in addition to other MCI subtypes with presumably different etiologies. Our relatively modest sample size precluded more thorough examination of variation and covariation among fluency measures in our individual participant groups.

4.4. Theoretical implications and directions for future research

Fluency tasks rely on numerous cognitive processes to a greater or lesser degree including accessing semantic memory stores and engaging various executive processes including initiation, monitoring, organization, rule implementation, and set shifting. Several models have been proposed to explain verbal fluency deficits in AD, and difficulties usually are ascribed to a differential loss or disorganization of semantic knowledge (Grober et al., 1985; Hodges et al., 1992; Monsch et al., 1992; Salmon et al., 1999), although some studies have highlighted impaired retrieval (Nebes, 1989; Nebes & Brady, 1988; Nebes et al., 1984). Most studies of AD that include both semantic and phonemic fluency tasks find a significantly larger semantic deficit thought to reflect the greater demands upon the integrity of the semantic network but the two measures are considered equivalent in sensitivity to executive control processes such as effortful retrieval (Henry et al., 2004). In our study, both phonemic and semantic fluency were reduced in MCI relative to cognitively intact older adults, suggesting subtle changes in both the quality of the semantic store and retrieval slowing. Further, switching ($\eta^2 = .245$) and shifting ($\eta^2 = .219$) fluencies showed the largest effect sizes, identifying them as most reliably discerning group membership, followed by category ($\eta^2 = .130$), letter fluency ($\eta^2 = .091$) and then action ($\eta^2 = .075$) (i.e., verb) generation and supermarket naming did not show group differences and appear to be less valuable clinically.

The present results suggest that tasks imposing substantial demands on both semantic stores and effortful retrieval will show greater effects due to the additive features of semantic knowledge degeneration and retrieval impairment. Correlational data further revealed that all fluency tasks correlated with WAIS-III Digit Symbol, a psychomotor processing speed task with known executive components. In addition, while category fluency correlated with tests known to assess semantic knowledge (i.e., WAIS-III Vocabulary and Information), switching fluency did not. Thus, although executive difficulties are known to be increasingly prominent in later neurodegenerative disease stages, the present results imply that the relative contributions of subtle executive versus semantic memory storage difficulties are largely equivalent at this early stage of disease. This finding may have implications for current conceptualizations of MCI (Winblad et al., 2004). Further investigation of the D-KEFS switching task is warranted to clarify the relative contributions of both semantic and executive processes to task performance.

4.5. Study limitations and conclusions

An important study limitation was the homogeneity of the participants with regard to baseline intellect, level of education, and ethnicity. Although reflective of the region of northern New England from which the sample
was drawn, these factors limit the generalizability of our findings, and we are actively recruiting a more demographically diverse cohort. Another limitation was the cross-sectional nature of our data, which prevented us from investigating the longitudinal course of verbal fluency performance and predictive value of fluency tasks. We are following-up with study participants and plan to investigate temporal changes in verbal fluency. Although adequate for the current set of analyses, a larger sample size would be desirable in addition to a more heterogeneous cohort of MCI participants (e.g., multiple-domain, single nonmemory domain) to enable further exploration of verbal fluency performance patterns and underlying cognitive processes tapped by specific tasks. Due to sample size constraints, we utilized data from the same group of participants for both the factor analysis and discriminant function analysis. Ideally, we would have used a validation sample in which factors derived from the first sample were used to assign classifications to an unrelated sample, and this represents a future direction for us.

Overall, study findings indicated that various subtypes of verbal fluency become statistically but not clinically compromised in the preclinical dementia stage of amnestic MCI. Factor analytic and correlational results further indicated involvement of both semantic stores and effortful retrieval in verbal fluency processes. The D-KEFS switching task most strongly discriminated group membership, and loaded on its own factor, suggesting that it measures somewhat different cognitive processes than other fluency tasks. These findings underscore the importance of including multiple verbal fluency tests (and other non-memory tasks) in assessment batteries targeting preclinical dementia populations. We also are examining relations of verbal fluency with structural neuroimaging data to investigate the neural basis for declining fluency abilities.

Acknowledgements

This research was supported by funding from the NIA (R01 AG19771), Alzheimer’s Association, National Science Foundation, and Ira DeCamp Foundation. Portions of this article were presented at the 25th Annual Conference of the National Academy of Neuropsychology in Tampa, FL, USA. We thank Margaret Nordstrom and Dr. Robert Santulli for their assistance with this project.

References


