Similar Verbal Fluency Patterns in Amnestic Mild Cognitive Impairment and Alzheimer’s Disease

Edmond Teng1,2,3,4, Judith Leone-Friedman3,4, Grace J. Lee5, Stephanie Woo6, Liana G. Apostolova3,4, Shelly Harrell6, John M. Ringman3,4, Po H. Lu4,∗

1Neuropsychology Unit, Veterans Affairs Greater Los Angeles Healthcare System, Los Angeles, CA, USA
2Geriatric Research Education and Clinical Center, Veterans Affairs Greater Los Angeles Healthcare System, Los Angeles, CA, USA
3Easton Center for Alzheimer’s Disease Research, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA
4Department of Neurology, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA
5Department of Psychology, Loma Linda University, Loma Linda, CA, USA
6Graduate School of Education and Psychology, Pepperdine University, Los Angeles, CA, USA

∗Corresponding author at: Department of Neurology, David Geffen School of Medicine at University of California Los Angeles, 710 Westwood Plaza., RNRC, Room 2-258, Los Angeles, CA 90095-1769, USA. Tel.: +1-310-794-3601; fax: +1-310-794-3148.
E-mail address: plu@mednet.ucla.edu (P.H. Lu).

Accepted 9 May 2013

Abstract

Disproportionately greater deficits in semantic relative to phonemic verbal fluency are seen in Alzheimer’s disease (AD) and have been attributed to neurodegenerative changes in the temporal lobe. Amnestic (AMN) mild cognitive impairment (MCI), which often represents incipient AD, is also characterized by early temporal lobe neuropathology, but previous comparisons of verbal fluency between AD and AMN MCI have yielded mixed results. We examined semantic and phonemic verbal fluency performance in 399 individuals (78 AD, 138 AMN MCI, 72 non-amnestic MCI, and 111 cognitively normal controls). Similar verbal fluency patterns were seen in AMN MCI and AD; both groups exhibited disproportionately poorer performance on semantic verbal fluency relative to normal controls. However, relative verbal fluency indices performed more poorly than individual semantic or phonemic verbal fluency indices for discriminating AMN MCI or AD participants from normal controls, suggesting that they are unlikely to provide additional utility for predicting progression from MCI to AD.

Keywords: Mild cognitive impairment; Alzheimer’s disease; Verbal fluency; Dementia; Assessment; Cognition

Introduction

Verbal fluency tests are widely used to assess language abilities and executive functioning. These tasks involve rapid associative exploration and retrieval of words based on phonemic (e.g., starting with a specified letter of the alphabet) or semantic/categorical (e.g., animals) criteria over a brief timed interval. Both tasks are thought to place comparable demands upon executive processes, requiring the efficient organization of verbal retrieval and recall, self-monitoring, effortful self-initiation, and inhibition of inappropriate responses (Henry & Crawford, 2004). However, although phonemic fluency necessitates search strategies based principally on broad lexical representations, semantic fluency requires a constrained search of items from a superordinate category and relies on semantic associations within the category itself. More words are typically generated on semantic verbal fluency tests than on phonemic verbal fluency tests because, in the latter paradigm, responses are not already grouped in organized presentations, and more restrictive rules, such as the exclusion of proper nouns, numbers, and morphological variants of words already generated, increase demands on monitoring and working memory and require greater executive control (Bialystok, 2011; Lezak, Howieson, & Loring, 2004).

Because both semantic and phonemic verbal fluency tasks are sensitive to executive functioning, frontal lobe lesions impair performance on both kinds of tests to a similar degree (Henry & Crawford, 2004). However, semantic verbal fluency tests have
an additional semantic memory component, and temporal lobe lesions result in relatively greater impairment in semantic verbal fluency when compared with phonemic verbal fluency (Henry & Crawford, 2004; Stuss et al., 1998). Patients with Alzheimer’s disease (AD) exhibit deficits in both semantic and phonemic verbal fluency, but prior meta-analyses have indicated that their impairments on semantic verbal fluency are disproportionately more severe (Henry, Crawford, & Phillips, 2004; Laws, Duncan, & Gale, 2010). This pattern remains stable with increasing dementia severity and may reflect both the important role of the temporal lobes in semantic memory and the relative predilection for neurodegenerative changes in the temporal lobe in AD (Henry et al., 2004), although alternative explanations have also been proposed. For instance, the discrepancies in semantic versus phonemic fluency seen in AD may also be present to a lesser degree in healthy controls and hence have been hypothesized to represent an exaggerated normal tendency (Laws et al., 2010).

Since neuropathological and neuroimaging changes associated with AD can be identified in the temporal lobe prior to a clinical diagnosis of AD (He et al., 2009; Kantarci et al., 2008; Riley, Snowdon, & Markesbery, 2002; Schneider, Arvanitakis, Leurgans, & Bennett, 2009), a number of studies have investigated semantic and phonemic fluency in mild cognitive impairment (MCI), particularly amnestic (AMN) MCI, which often represents the prodromal stage of incipient AD (Bruscoli & Lovestone, 2004; Petersen, 2011). Although these studies have been motivated by the potential for using verbal fluency tests, which are easily administered in clinical settings, to improve the prediction of subsequent progression from MCI to AD, they have produced mixed results. Some investigators have reported that participants with MCI perform similar to AD patients, demonstrating relatively poorer semantic versus phonemic verbal fluency when compared with cognitively normal elderly controls (Adlam, Bozeat, Arnold, Watson, & Hodges, 2006; Cottingham & Hawkins, 2010; Lonie et al., 2009; Murphy, Rich, & Troyer, 2006), whereas others have reported similar patterns of verbal fluency in MCI and control cohorts (Brandt & Manning, 2009; Nutter-Upham et al., 2008). Although the underlying reasons for these discrepancies remain uncertain, Brandt and Manning (2009) have postulated that they may arise from differences in the psychometric properties of specific verbal fluency assessments and/or the operationalization of AMN MCI between research centers. The relatively small study cohorts included in these earlier reports, with sample sizes ranging from 10 to 74 MCI participants, 11 to 35 AD patients, and 24 to 46 healthy controls (Adlam et al., 2006; Brandt & Manning, 2009; Lonie et al., 2009; Murphy et al., 2006; Nutter-Upham et al., 2008), may further limit their interpretation.

The main objective of the current study was to clarify the utility of comparing relative performance on semantic versus phonemic verbal fluency for distinguishing between larger cohorts of cognitively normal, MCI, and mild AD participants. Our study extends and contributes to the existing literature in several important ways. First, we sought to replicate previous findings of category and phonemic verbal fluency deficits in MCI and AD with a substantially larger cohort of well-characterized participants. Second, prior studies have focused predominantly on AMN MCI with limited data on non-amnestic (NON) subtypes, even though the latter group also demonstrates elevated rates of progression to AD relative to cognitively normal controls (Rountree et al., 2007). Our MCI cohort has been further classified into AMN (both the single domain and the multiple domain) and NON subgroups in order to investigate the pattern of verbal fluency performance across different MCI subtypes and examine which MCI subgroups would exhibit verbal fluency profiles most similar to those seen in the mild AD group. Our hypothesis was that MCI subgroups associated with the highest rates of progression to AD (i.e., AMN multiple domain [mdAMN]; Busse, Hensel, Guhne, Angermeyer, & Riedel-Heller, 2006; Fischer et al., 2007; Nordlund et al., 2010; Rasquin, Lodder, Visser, Lousberg, & Verhey, 2005; Yaffe, Petersen, Lindquist, Kramer, & Miller, 2006) would exhibit patterns of verbal fluency performance most similar to those seen in the mild AD group. Third, we examined both raw difference scores, which are easier to apply in clinical practice, and standardized difference scores, which account for the inherent differences in difficulty between individual semantic and phonemic verbal fluency tasks and the influences of age and education on these tasks. Finally, we also used receiver operating characteristic (ROC) curves to determine the relative utility of conventional verbal fluency scores versus verbal fluency difference scores for cross-sectional diagnostic classifications.

Methods

Participants

Participants were drawn from a larger research cohort enrolled in an ongoing study at the Easton Center for Alzheimer’s Disease Research at UCLA and were initially recruited from patients seen in the UCLA Memory Disorders clinic or from community volunteers. Research assessments included neuropsychological testing, physician interview, and neurological examination. The neuropsychological battery evaluated memory, attention, language, visuospatial, and executive function as described previously (Teng, Lu, & Cummings, 2007). The specific tests incorporated in the battery are listed in Supplementary material online, Table S1. Interviews with participants and informants included questions regarding the performance of basic and instrumental activities of daily living (ADLs). Research diagnoses were determined in a multidisciplinary conference, based on neuropsychological testing scores and confirmed by consensus clinician judgment regarding ADL performance.
Participants in the AD group fulfilled the National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer’s Disease and Related Disorders Association (NINCDS/ADRDA) criteria (McKhann et al., 1984) for possible or probable AD. Participants in the MCI subgroups fulfilled modified Petersen criteria (Petersen, 2004): (a) subjective cognitive complaint, (b) essentially intact ADLs, (c) objective cognitive impairment, and (d) not demented. Objective cognitive impairment was operationalized as performance $\geq 1.5$ SD below published age and education adjusted normative means on at least one neuropsychological test in any domain. Participants meeting criteria for MCI were subdivided into AMN or NON groups based on the presence or the absence of impairment on one or more of the memory tests included in our neuropsychological battery (Supplementary material online, Table S1). The AMN subgroup was further classified as single-domain (sdAMN; impairment limited to the memory domain) or mdAMN (impairments in memory and at least one other domain). The NON group was not further subdivided in this fashion due to the relatively small number of subjects meeting criteria for multiple-domain NON MCI ($n = 14$). The normal comparison (NC) group consisted of individuals who performed within the normal range on all cognitive assessments, irrespective of subjective cognitive complaints. Exclusion criteria included: (a) age $< 50$, (b) diagnosis of dementia syndrome other than AD, (c) MRI or CT of the brain demonstrating any major focal lesions (mild to moderate microvascular ischemic changes or isolated lacunes noted on clinical neuroradiology reports were permitted), and (d) Mini-Mental Status Examination (MMSE; Folstein, Folstein, & McHugh, 1975) scores $< 20$.

Verbal Fluency Indices

Two measures of verbal fluency were administered to each participant as part of the larger neuropsychological battery. Semantic verbal fluency was assessed by asking participants to generate as many animals as possible within 60 s (Benton & Hamsher, 1989). The Controlled Oral Word Association Test (COWAT; Benton & Hamsher, 1989) was used to assess phonemic verbal fluency. Participants were asked to generate as many words as possible beginning with the letters “F,” “A,” and “S”; 60 s were given for each letter. Several word-generation rules were also explained to the participants: (a) no word could be a proper noun (i.e., the name of a person or a place such as “Bob” or “Boston”) and (b) no word could be simply repeated with a different suffix (i.e., “big,” “bigger,” and “biggest”).

The total number of unique words generated was tabulated. Relative performances on categorical versus phonemic verbal fluency were compared between diagnostic groups using two different analyses: (a) differences in raw scores and (b) differences in standardized scores calculated from published normative data (Tombaugh, Kozak, & Rees, 1999).

Data Analyses

Statistical analyses were performed using SPSS 16.0 for Mac (SPSS Inc., Chicago, IL, USA). Demographic data were compared between groups using one-way analysis of variance (ANOVA) or unpaired $t$-tests for continuous variables and Kruskal–Wallis or chi-squared tests for categorical variables. Verbal fluency indices were compared with unpaired $t$-tests, one-way ANOVA, or analysis of co-variance (ANCOVA) where appropriate. Post hoc analyses were Bonferroni corrected for multiple comparisons (critical $p = .008$). The utility of semantic verbal fluency scores, phonemic verbal fluency scores, and semantic–phonemic difference scores for group classification was assessed using area under the curve (AUC) analyses of ROC curves.

Table 1. Demographic data for the NC, NON, AMN, and AD groups

<table>
<thead>
<tr>
<th></th>
<th>NC</th>
<th>NON</th>
<th>AMN</th>
<th>AD</th>
<th>$F$/$\chi^2$ (3,395)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>111</td>
<td>72</td>
<td>138</td>
<td>78</td>
<td>—</td>
</tr>
<tr>
<td>Age</td>
<td>70.3 (8.0)</td>
<td>72.3 (8.5)</td>
<td>73.8 (8.2)$^a$</td>
<td>76.3 (7.9)$^{a,b}$</td>
<td>8.86$^*$</td>
</tr>
<tr>
<td>Education</td>
<td>17.0 (2.3)</td>
<td>16.4 (2.7)</td>
<td>15.6 (2.7)$^a$</td>
<td>15.2 (2.8)$^{a,b}$</td>
<td>8.85$^*$</td>
</tr>
<tr>
<td>% Male</td>
<td>55.9%</td>
<td>58.3%</td>
<td>49.3%</td>
<td>50.0%</td>
<td>2.23</td>
</tr>
<tr>
<td>% Non-Hispanic White</td>
<td>94.6%</td>
<td>86.1%</td>
<td>86.2%</td>
<td>91.0%</td>
<td>5.68</td>
</tr>
<tr>
<td>MMSE</td>
<td>29.1 (1.2)</td>
<td>28.1 (1.9)$^a$</td>
<td>26.9 (2.2)$^{a,b}$</td>
<td>23.5 (2.4)$^{a,b,c}$</td>
<td>131.1$^*$</td>
</tr>
</tbody>
</table>

Notes: NC = normal control; NON = non-amnestic MCI; AMN = amnestic MCI; AD = Alzheimer’s disease. Parentheses indicate the standard deviation. Post-hoc analyses are adjusted for multiple comparisons with Bonferroni correction ($p < .008$ vs. NC$^a$, NON$^b$, or AMN$^c$). $^a p < .05$ for overall ANOVA or Kruskal–Wallis test.
Results

Demographics

Demographic data for the NC, AD, AMN, and NON groups are shown in Table 1. Similar gender and ethnic distributions were seen across groups. Significant group differences were seen with age, education, and MMSE scores. Post hoc analyses indicated that the AMN group was significantly older than the NC group \((p = .001)\) and the AD group was significantly older than both the NC \((p < .001)\) and NON \((p = .003)\) groups. The AMN group averaged fewer years of education than the NC group \((p < .001)\), and the AD group averaged fewer years of education than the NC \((p < .001)\) and NON \((p = .006)\) groups. MMSE scores differed significantly between all groups \((NC > NON > AMN > AD; all ps \leq .002)\). Age and years of education have previously been reported to be significant mediators of verbal fluency performance \((\text{Mathuranath et al., } 2003; \text{Rodriguez-Aranda & Martinussen, } 2006; \text{Tombaugh et al., } 1999)\). Therefore, subsequent analyses of raw verbal fluency scores were adjusted for these demographic variables.

Raw Verbal Fluency Scores

Raw verbal fluency data for each experimental group are shown in Fig. 1 and Supplementary material online, Table S2. ANCOVAs that included age and education as covariates revealed significant group differences on overall Animals, \(F(3, 393) = 59.87, p < .001\), and COWAT, \(F(3, 393) = 33.67, p < .001\), scores, as well as individual letter scores for F, \(F(3, 383) = 22.96, p < .001\), A, \(F(3, 383) = 33.14, p < .001\), and S, \(F(3, 383) = 24.72, p < .001\). Slightly fewer degrees of freedom were available for individual COWAT letter analyses because of missing individual letter data for 10 participants \((3 NC, 3 NON, 2 AMN, and 2 AD)\). Post hoc analyses for each of these raw verbal fluency indices indicated that the NC group performed better than the NON, AMN, and AD groups \((all ps \leq .002)\) and the NON and AMN groups performed better than the AD group \((all ps < .001)\).

For each participant, difference scores were calculated between the raw semantic and phonemic verbal fluency scores to determine relative performance on these indices. Phonemic verbal fluency (average FAS letter \([i.e., \text{total COWAT score}/3]\)) and individual “F,” “A,” and “S”) scores were subtracted from categorical verbal fluency (i.e., Animals) scores to generate difference scores. ANCOVAs using age and education as covariates revealed significant group effects on the average FAS letter difference score \((Fig. 2A and Supplementary material online, Table S2), F(3, 393) = 7.24, p < .001\), as well as on each of the individual letter difference scores \((Fig. 2B and Supplementary material online, Table S2): F, F(3, 383) = 9.34, p < .001; A, F(3, 383) = 5.19, p = .002; S, F(3, 383) = 5.30, p = .001\). Post hoc analyses of the Animals-average FAS letter and Animals-“F” measures indicated that both the AMN and AD groups had lower scores than the NC group \((ps \leq .002)\) and that the AD group had lower scores than the NON group \((ps \leq .002)\). On the Animals-“A” measure, the AD group had lower scores than both the NC and NON groups \((ps \leq .004)\), and on the Animals-“S” measure, both the AMN and AD groups had lower scores than the NC group \((ps \leq .002)\).

![Fig. 1](image_url) Raw verbal fluency scores for (A) Animals, (B) total COWAT, and (C) individual COWAT letters “F,” “A,” and “S” in the NC, NON MCI, AMN MCI, and AD groups. Error bars represent the standard error of the mean. Post hoc analyses were Bonferroni corrected \((p < .008 \text{ vs. } \ast NC, \text{†NON, or ‡AMN groups})\).
Standardized Verbal Fluency Scores

Difference scores calculated from raw verbal fluency indices may be affected by the total number of words produced and by the relative difficulty of individual tests (Brandt & Manning, 2009). An alternative approach is to examine standardized rather than raw verbal fluency scores. Published normative verbal fluency data (Tombaugh et al., 1999) were used to calculate standardized Animals and COWAT z-scores (Fig. 3 and Supplementary material online, Table S2). One-way ANOVAs indicated significant group effects on both the Animals, \( F(3, 395) = 72.65, p < .001 \), and COWAT, \( F(3, 395) = 34.39, p < .001 \), z-scores. Post hoc analyses indicated that the NC group performed significantly better on both standardized verbal fluency indices than each of the other three groups (all \( ps < .001 \)). On the Animals z-score, the NON group performed better than the AMN and AD groups (\( ps < .004 \)) and the AMN group performed better than the AD group (\( p < .001 \)). On the COWAT z-score, both the NON and AMN groups performed better than the AD group (\( ps < .001 \)).

Semantic–phonemic difference scores were also calculated using standardized z-scores (Fig. 4 and Supplementary material online, Table S2). There was again an overall effect of group, \( F(3, 395) = 9.24, p < .001 \). Post hoc analyses revealed significantly greater standardized difference scores in the NC group relative to the AMN and AD groups (\( ps < .001 \)) and in the NON group relative to the AD group (\( p < .001 \)).

Verbal Fluency Scores in sdAMN Versus mdAMN MCI

Both the AMN and AD groups demonstrated disproportionally greater deficits on semantic verbal fluency than the NC and NON groups. Since several studies have reported higher rates for progression to AD in mdAMN MCI than in sdAMN MCI (Busse et al., 2006; Manly et al., 2008; Rasquin et al., 2005; Tabert et al., 2006), we sought to examine whether the corresponding subgroups in our cohort (sdAMN and mdAMN) demonstrated different patterns of performance on categorical versus phonemic verbal fluency as has been reported previously (Brandt & Manning, 2009).
Demographic data for the sdAMN and mdAMN groups are shown in Table 2. The mdAMN group was significantly younger than the sdAMN group \((p = .024)\), had a marginally lower percentage of non-Hispanic Whites \((p = .053)\), and had marginally lower MMSE scores \((p = .090)\). Of the 89 participants included in the mdAMN group, 46 had deficits in one additional domain, 19 had deficits in two additional domains, 17 had deficits in three additional domains, and 7 had deficits in four additional domains. Across the entire mdAMN group, 30.3% had attention/processing speed deficits, 42.7% had language deficits, 40.4% had visuospatial deficits, and 69.7% had frontal/executive deficits.

Raw verbal fluency data for these two subgroups are shown in Fig. 5 and Supplementary material online, Table S3. Individual COWAT letter data were missing for two participants in the mdAMN group. ANCOVAs adjusted for age and education indicated that the mdAMN group performed significantly more poorly on Animals, \(F(1, 134) = 17.67, p < .001\), overall COWAT, \(F(1, 134) = 11.21, p = .001\), and each of the individual COWAT letter analyses—F, \(F(1, 132) = 6.09, p = .015\); A, \(F(1, 132) = 12.28, p = .001\); S, \(F(1, 132) = 8.35, p = .005\). Despite these group differences on categorical and phonemic verbal fluency, there were no differences between the sdAMN and mdAMN groups on raw or standardized difference scores (Supplementary material online, Table S3).

Verbal Fluency Scores for Group Prediction

Both the AD and AMN groups demonstrated disproportionately greater deficits on semantic relative to phonemic verbal fluency. Therefore, we further sought to determine whether calculated raw and/or standardized verbal fluency difference scores yielded better predictive utility than individual semantic or phonemic fluency scores alone for distinguishing these two groups from the NC group. ROC curves comparing the relative predictive power of these verbal fluency indices are shown in Fig. 6. Each of these measures performed significantly better than chance (all \(p < .003\)) for distinguishing between the NC and AD groups (Fig. 6A and B) and between the NC and AMN groups (Fig. 6C and D). However, the highest AUC values were obtained with the semantic verbal fluency scores alone, and the AUC values for the raw and standardized semantic–phonemic verbal fluency difference scores were lower than those calculated for individual semantic or phonemic verbal fluency.

Table 2. Demographic data for the sdAMN and mdAMN MCI subgroups (parentheses indicate the standard deviation)

<table>
<thead>
<tr>
<th></th>
<th>sdAMN</th>
<th>mdAMN</th>
<th>(t^2(146))</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>49</td>
<td>89</td>
<td>—</td>
</tr>
<tr>
<td>Age</td>
<td>75.9 (8.4)</td>
<td>72.6 (7.8)</td>
<td>2.28*</td>
</tr>
<tr>
<td>Education</td>
<td>16.1 (2.6)</td>
<td>15.4 (2.7)</td>
<td>1.50</td>
</tr>
<tr>
<td>% Male</td>
<td>53.1%</td>
<td>47.2%</td>
<td>0.44</td>
</tr>
<tr>
<td>% Non-Hispanic White</td>
<td>93.9%</td>
<td>82.0%</td>
<td>3.74</td>
</tr>
<tr>
<td>MMSE</td>
<td>27.3 (2.1)</td>
<td>26.7 (2.3)</td>
<td>1.71</td>
</tr>
</tbody>
</table>

Notes: sdAMN = single-domain amnestic; mdAMN = multiple-domain amnestic; MMSE = Mini-Mental State Examination.

*\(p < .05\).
Our results support the hypothesis that individuals meeting criteria for AMN MCI demonstrate greater deficits on semantic versus phonemic fluency compared with cognitively normal elderly controls, who tend to perform better on semantic than phonemic fluency tasks (Gladsjo et al., 1999; Tombaugh et al., 1999). These findings are consistent with several prior studies of comparable cohorts (Adlam et al., 2006; Cottingham & Hawkins, 2010; Lonie et al., 2009; Murphy et al., 2006). The relatively greater impairment on semantic verbal fluency seen in our AMN group was similar in pattern to that seen in our mild AD group, albeit less pronounced. Although our NON group also demonstrated impairments on both semantic and phonemic fluency tasks, their overall pattern of verbal fluency performance was more similar to that seen in our NC group (i.e., relatively better performance on semantic vs. phonemic verbal fluency).

![Fig. 5](image1.png)

**Fig. 5.** Raw verbal fluency scores for (A) Animals, (B) total COWAT, and (C) individual COWAT letters in the sdAMN and mdAMN MCI groups. Error bars represent the standard error of the mean. *p < .05 versus sdAMN.

![Fig. 6](image2.png)

**Fig. 6.** ROC curves of raw (A and C) and standardized (B and D) verbal fluency indices for distinguishing NC and AD (A and B) and NC and AMN (C and D) groups. AUC indicates the area under the curve.

**Discussion**

Our results support the hypothesis that individuals meeting criteria for AMN MCI demonstrate greater deficits on semantic versus phonemic fluency compared with cognitively normal elderly controls, who tend to perform better on semantic than phonemic fluency tasks (Gladsjo et al., 1999; Tombaugh et al., 1999). These findings are consistent with several prior studies of comparable cohorts (Adlam et al., 2006; Cottingham & Hawkins, 2010; Lonie et al., 2009; Murphy et al., 2006). The relatively greater impairment on semantic verbal fluency seen in our AMN group was similar in pattern to that seen in our mild AD group, albeit less pronounced. Although our NON group also demonstrated impairments on both semantic and phonemic fluency tasks, their overall pattern of verbal fluency performance was more similar to that seen in our NC group (i.e., relatively better performance on semantic vs. phonemic verbal fluency).
It has been suggested that the relatively poorer performance on semantic verbal fluency tasks in AD is due to the disproportionately greater burden of neurodegenerative changes in the temporal lobe, particularly earlier in the disease course (Henry et al., 2004). This hypothesis is supported by analyses of patients with focal lesions, which have demonstrated that disproportionate deficits in semantic verbal fluency are associated with temporal lobe lesions (Baldo, Schwartz, Wilkins, & Drorkers, 2006; Henry & Crawford, 2004). Furthermore, structural imaging analyses in MCI and AD indicate that semantic verbal fluency deficits in these groups correlate with decreased gray matter density in the left hippocampus and the temporal cortex (Dos Santos et al., 2011).

The relatively greater semantic versus phonemic verbal fluency deficits seen in AMN MCI but not in NON MCI in the current study likely reflect different patterns of underlying neuropathology in these two groups, particularly the relative predominance of early temporal lobe neurodegeneration in AMN MCI. Structural magnetic resonance imaging studies have indicated significant hippocampal atrophy relative to cognitive normal elderly controls in AMN MCI but not NON MCI (He et al., 2009; Kantarci et al., 2008). Likewise, autopsy studies report that AMN MCI cases have greater limbic neurofibrillary tangle involvement (Riley et al., 2002) and a higher likelihood of fulfilling neuropathological criteria for AD (Schneider et al., 2009) than NON MCI cases. Furthermore, most (Busse et al., 2006; Fischer et al., 2007; Nordlund et al., 2010; Rasquin et al., 2005; Yaffe et al., 2006), though not all (Rountree et al., 2007), longitudinal studies of MCI have demonstrated that AMN deficits are associated with higher rates of progression to AD. The discrepant findings may be attributed to methodological differences between studies involving study sample composition, operational definitions of MCI, neuropsychological instruments, and duration of follow-up, but most studies suggest that AMN MCI is more likely to represent more advanced incipient AD than NON MCI.

Our results diverge to some extent from two prior studies of verbal fluency in AMN MCI. Nutter-Upham and colleagues (2008) reported similar deficits on semantic and phonemic verbal fluency in their AMN MCI cohort. However, although the raw difference scores in their cognitively normal and AMN MCI groups were similar in overall magnitude, the effect size for the AMN MCI group’s deficit on semantic verbal fluency (Cohen’s $d = 0.98$) was larger than the effect size for their deficit on phonemic verbal fluency (Cohen’s $d = 0.66$), due to greater variability in phonemic verbal fluency performance. Therefore, an alternative interpretation is that their findings actually do reflect a preferential deficit on semantic verbal fluency in AMN MCI.

Brandt and Manning (2009) reported that the pattern of semantic versus phonemic verbal fluency in their sdAMN MCI cohort was more similar to their cognitively normal elderly control group, whereas their mdAMN MCI group performed more similar to their mild to moderate AD group. Although our mdAMN group performed more poorly on both semantic and phonemic verbal fluency than our sdAMN group, the overall pattern of semantic versus phonemic verbal fluency deficits, as represented by raw or standardized difference scores, was similar between the two groups. The definitive explanation for these conflicting findings remains uncertain but may be related to differences between studies in cohort size or the operationalization of sdAMN MCI. Specifically, the Brandt and Manning (2009) psychometric definition for sdAMN MCI involved performance $\geq 1.5 SD$ below normative means on WMS-R Logical Memory delayed recall, while we operationalized our sdAMN participants as performing $\geq 1.5 SD$ below published normative means on any one of the four memory tasks included in our battery (Supplementary material online, Table S1). The slightly broader neuropsychological criteria used in the current study might have been expected to produce a higher false-positive rate for identifying incipient AD (Teng, Tingus, Lu, & Cummings, 2009). Brandt and Manning (2009) also suggested that previous reports of disproportionate deficits on semantic verbal fluency in AMN MCI might be driven more by individual task difficulty rather than differences related to semantic versus phonemic demands. However, in the current study, we saw similar verbal fluency performance patterns using three phonemic cues with different levels of difficulty, and the disproportionate deficits in semantic verbal fluency in the AMN group remained robust with either raw or standardized difference scores.

Given the disproportionate deficits in semantic versus phonemic verbal fluency in our AMN and AD groups, we further sought to determine whether the difference scores used to quantify these discrepancies would have further utility for the prediction of group membership, and by extension, potential utility for identifying AMN MCI individuals at highest risk for subsequent progression to AD. These analyses were motivated by previous work suggesting the use of verbal fluency difference scores for such purposes (Lonie et al., 2009). However, our ROC analyses indicated that both raw and standardized verbal fluency difference scores yielded only mediocre discrimination between the NC group and either the AD or the AMN group. Indeed, group membership predictions using the composite verbal fluency measures performed more poorly than predictions based on either semantic or phonemic verbal fluency alone. These results indicate that although there is an overall group effect toward relatively poorer semantic verbal fluency in the AMN and AD groups, the substantial overlap in difference scores between memory impaired and cognitively intact groups may preclude predictive utility for individual patients in clinical settings.

There are a number of factors that may limit the interpretation of our results. Our overall study cohort was a highly educated convenience sample predominantly comprised of non-Hispanic White participants. In particular, our NC group performed better than expected on both semantic and phonemic verbal fluency relative to published normative data (Tombaugh et al., 1999). Therefore, our results may be less generalizable to study cohorts with lower levels of formal education or more diverse ethnic backgrounds. There were significant demographic differences between diagnostic groups, including in age and education, factors that have been shown to impact verbal fluency performance (Mathuranath et al., 2003; Rodríguez-Aranda & Martinussen, 2002) psychometric definition for sdAMN MCI involved performance $\geq 1.5 SD$ below normative means on WMS-R Logical Memory delayed recall, while we operationalized our sdAMN participants as performing $\geq 1.5 SD$ below published normative means on any one of the four memory tasks included in our battery (Supplementary material online, Table S1). The slightly broader neuropsychological criteria used in the current study might have been expected to produce a higher false-positive rate for identifying incipient AD (Teng, Tingus, Lu, & Cummings, 2009). Brandt and Manning (2009) also suggested that previous reports of disproportionate deficits on semantic verbal fluency in AMN MCI might be driven more by individual task difficulty rather than differences related to semantic versus phonemic demands. However, in the current study, we saw similar verbal fluency performance patterns using three phonemic cues with different levels of difficulty, and the disproportionate deficits in semantic verbal fluency in the AMN group remained robust with either raw or standardized difference scores.

Given the disproportionate deficits in semantic versus phonemic verbal fluency in our AMN and AD groups, we further sought to determine whether the difference scores used to quantify these discrepancies would have further utility for the prediction of group membership, and by extension, potential utility for identifying AMN MCI individuals at highest risk for subsequent progression to AD. These analyses were motivated by previous work suggesting the use of verbal fluency difference scores for such purposes (Lonie et al., 2009). However, our ROC analyses indicated that both raw and standardized verbal fluency difference scores yielded only mediocre discrimination between the NC group and either the AD or the AMN group. Indeed, group membership predictions using the composite verbal fluency measures performed more poorly than predictions based on either semantic or phonemic verbal fluency alone. These results indicate that although there is an overall group effect toward relatively poorer semantic verbal fluency in the AMN and AD groups, the substantial overlap in difference scores between memory impaired and cognitively intact groups may preclude predictive utility for individual patients in clinical settings.

There are a number of factors that may limit the interpretation of our results. Our overall study cohort was a highly educated convenience sample predominantly comprised of non-Hispanic White participants. In particular, our NC group performed better than expected on both semantic and phonemic verbal fluency relative to published normative data (Tombaugh et al., 1999). Therefore, our results may be less generalizable to study cohorts with lower levels of formal education or more diverse ethnic backgrounds. There were significant demographic differences between diagnostic groups, including in age and education, factors that have been shown to impact verbal fluency performance (Mathuranath et al., 2003; Rodríguez-Aranda & Martinussen, 2002).
However, our results remained robust after statistical adjustments for these differences. The absence of recognition memory components for three of the four memory tasks in our battery precluded the inclusion of delayed recognition scores for clinical diagnoses of AMN MCI, but future studies should examine discrepancies between recall and recognition memory, which may increase the specificity for diagnosing AMN MCI due to AD. Clinical diagnoses of NON MCI (but not AMN MCI) incorporated verbal fluency performance, which may have skewed the results of the NON group. Nevertheless, given the small proportions of participants in the NON group with impaired performance on Animals (12.5%) or the COWAT (11.1%), this is unlikely to have significantly affected our results. Indeed, the relative patterns of verbal fluency performance in the NON group did not differ significantly from the NC group, regardless of whether the analyses included or excluded participants with deficits on Animals or the COWAT. The majority of our NON group (81%) was comprised of participants with single-domain NON deficits. The high reversion rates of single-domain NON MCI to normal cognition (Busse et al., 2006) and relatively poor longitudinal stability of NON deficits in MCI (Teng et al., 2009) raise the possibility that our NON group is less likely to have underlying neurodegenerative disease, which may represent an alternative explanation for their similarity with the NC group on relative verbal fluency measures. Furthermore, healthy adults may perform poorly on one or two tasks when administered a large battery of tests (Binder, Iverson, & Brooks, 2009; Palmer, Boone, Lesser, & Wohl, 1998; Schretlen, Munro, Anthony, & Pearlson, 2003). Therefore, the NON group may include a small proportion of healthy controls, which may further contribute to the similarity in relative verbal fluency scores between the NON and NC groups. Although we used multiple measures of phonemic verbal fluency, we used only a single measure of semantic verbal fluency, as a significant subset of our participants were assessed before a second semantic verbal fluency task was added to our neuropsychological testing battery. Brandt and Manning (2009) have advocated for the use of multiple measures for determining relative performance on semantic versus phonemic verbal fluency because of possible differences in difficulty across individual tasks. Nevertheless, the consistency of the results incorporating the three individual phonemic verbal fluency tests and the concordance of findings between raw and standardized difference scores, the latter of which provides a control for task difficulty, appear to support the reliability of our conclusions.

The results of the current cross-sectional study partially reconcile the discordant findings from previous studies of semantic versus phonemic verbal fluency in MCI cohorts. Our findings confirm prior work suggesting disproportionate deficits in semantic verbal fluency in AMN MCI (Adlam et al., 2006; Cottingham & Hawkins, 2010; Lonie et al., 2009; Murphy et al., 2006) but also emphasize that relative indices of semantic versus phonemic fluency are less likely to have significant incremental utility for predicting subsequent progression from MCI to AD on an individual basis (Brandt & Manning, 2009). Further longitudinal studies will be required to determine if relative verbal fluency indices may have greater utility for determining MCI progression when used in conjunction with other potential predictor variables. Future studies should also incorporate item-level responses to examine cluster and switching strategies during verbal fluency performance, which can potentially differentiate between executive and semantic knowledge deficits (Haugrud, Crossley, & Vrbancic, 2011; Price et al., 2012).

**Supplementary material**

Supplementary material is available at *Archives of Clinical Neuropsychology* online.

**Funding**

This research was supported by NIH grants from the National Institute on Aging (P50 AG 16570, K23 AG 028727 to PL and K08 AG 34628 to ET; jointly sponsored by the National Institute on Aging, American Federation for Aging Research, the John A. Hartford Foundation, the Atlantic Philanthropies, the Starr Foundation and an anonymous donor), the Alzheimer’s Disease Research Centers of California, the Sidell-Kagan Foundation, and the Jim Easton Consortium for Alzheimer’s Disease Drug Discovery and Biomarkers at UCLA.

**Conflict of Interest**

None declared.

**Acknowledgements**

We would like to thank Dahyun Yi and Christine Kang for their assistance with data collection and management.


