



Published in final edited form as:

*Neuropsychologia*. 2015 July ; 73: 169–175. doi:10.1016/j.neuropsychologia.2015.04.034.

## Free and Cued Memory in relation to Biomarker-Defined Abnormalities in Clinically Normal Older Adults and Those at Risk for Alzheimer's Disease

Kathryn V. Papp<sup>1</sup>, Rebecca E. Amariglio<sup>1</sup>, Elizabeth Mormino<sup>2</sup>, Trey Hedden<sup>3,4</sup>, Maria Dekhytar<sup>1</sup>, Keith A. Johnson<sup>2,3,4</sup>, Reisa A. Sperling<sup>1,2,3,4</sup>, and Dorene M. Rentz<sup>1,2</sup>

<sup>1</sup>Center for Alzheimer Research and Treatment, Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115

<sup>2</sup>Department of Neurology, Massachusetts General Hospital, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114

<sup>3</sup>Department of Radiology, Massachusetts General Hospital, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114

<sup>4</sup>Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, Massachusetts 02129

### Abstract

**Objectives**—Furthering our understanding of the relationship between amyloidosis (A $\beta$ ), neurodegeneration (ND), and cognition is imperative for early identification and early intervention of Alzheimer's disease (AD). However, the subtle cognitive decline differentially associated with each biomarker-defined stage of preclinical AD has yet to be fully characterized. Recent work indicates that different components of memory performance (free and cued recall) may be differentially specific to memory decline in prodromal AD. We sought to examine the relationship between free and cued recall paradigms, in addition to global composites of memory, executive functioning, and processing speed in relation to stages of preclinical AD.

**Methods**—A total of 260 clinically normal (CN) older adults (CDR=0) from the Harvard Aging Brain study were grouped according to preclinical AD stages including Stage 0 (A $\beta$ -/ND-), Stage

---

Corresponding Author: Kathryn V. Papp, kpapp@partners.org, Address: Center for Alzheimer's Research and Treatment, Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, 221 Longwood Ave., Boston, MA 02115, Telephone: (617) 486-5322, Fax number: (617) 726-5760.

**DISCLOSURES:** KVP and REA are co-investigators for Eisai, Eli Lilly, and Merck. DMR is a co-investigator for Eli Lilly. KAJ has served as paid consultant for Bayer, Biogen Idec, Bristol-Myers Squibb, GE Healthcare, Isis Pharmaceuticals Inc, Janssen Alzheimer's Immunotherapy, Piramal, Siemens Medical Solutions, and Genzyme. He is a site principal investigator co-investigator for Lilly/Avid, Biogen Idec, Bristol-Myers Squibb, Eisai, Pfizer, Janssen Alzheimer Immunotherapy, Merck, and Navidea clinical trials. RAS has served as a paid consultant for Bristol-Myers Squibb, Eisai, Janssen Alzheimer Immunotherapy, Pfizer, Merck, and Roche, and as an unpaid consultant to Avid, Eli Lilly. She is a site principal investigator co-investigator for Avid, Bristol-Myers Squibb, Pfizer, and Janssen Alzheimer Immunotherapy clinical trials. None of the above relationships are related to the content in the manuscript.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1 (A $\beta$ +/ND-), Stage 2 (A $\beta$ +/ND+), and suspected non-Alzheimer's associated pathology (SNAP; A $\beta$ -/ND+). General linear models controlling for age, sex, and education were used to assess for stage-based performance differences on cognitive composites of executive functioning, processing speed, and memory in addition to free and cued delayed recall on the Selective Reminding Test (SRT) and Memory Capacity Test (MCT).

**Results**—Global memory performance differed between preclinical stages with Stage 2 performing worse compared with Stage 0. When examining free and cued paradigms by memory test, only the MCT (and not the SRT) revealed group differences. More specifically, Stage 1 was associated with decrements in free recall compared with Stage 0 while Stage 2 was associated with decrements in both free and cued recall. There was a trend for the SNAP group to perform worse on free recall compared with Stage 0. Finally, there was no association between preclinical stage and global composites of executive functioning or processing speed.

**Conclusions**—Clinically normal older adults with underlying evidence of amyloidosis and neurodegeneration exhibit subtle, yet measurable differences in memory performance, but only on a challenging associative test. The sensitivity of free vs. cued memory paradigms may be dependent on preclinical stage such that reduced free recall is associated with amyloidosis alone (Stage 1) while a decline in cued recall may represent progression to amyloidosis and neurodegeneration (Stage 2). These findings may have practical applications for clinical assessment and clinical trial design.

## Keywords

free and cued memory; biomarker stages; preclinical Alzheimer's disease

## 1. INTRODUCTION

The recently developed (Sperling et al., 2011a) and operationalized (Jack et al., 2012, Knopman et al., 2012, Mormino et al., 2014a) biomarker-defined stages of preclinical Alzheimer's Disease (AD) have provided a common rubric for studying early disease: where individuals are defined as clinically normal but exhibit amyloidosis (A $\beta$ +), neurodegeneration (ND), and/or subtle cognitive decline. While only the final preclinical stage is defined as involving cognitive decline, we were interested in determining associations between all biomarker stages and cognition. Clinical trials for disease-modifying therapies are currently underway in older individuals in preclinical AD with the goal of preventing cognitive decline (Sperling et al., 2011b). However, whether subtle cognitive decline is associated with preclinical biomarker stages has yet to be fully characterized, and traditional neuropsychological measures may be insufficiently specific or challenging enough for the purpose of early detection and tracking (Rentz et al., 2013). It is important to know which cognitive measures might be most sensitive for selecting potential subjects to enroll in prevention trials prior to committing to expensive imaging or lumbar puncture procedures.

Tests of associative binding, such as the Free and Cued Selective Reminding Test, (Grober et al., 2000) and the Memory Capacity Test (MCT), (H. Buschke, personal communication, Rentz et al., 2010) show promise in distinguishing older adults in the preclinical phase of

AD (Rentz et al., 2013, Dubois et al., 2010). These measures improve encoding specificity by pairing the word to be remembered with a semantic cue, inducing deep encoding to maximize learning and recall (Grober et al., 2010; Rentz et al., 2013). Recall deficits that do not improve with cueing may be particularly indicative of temporolimbic amnesia (Dubois et al., 2010, Sarazin et al., 2007). A recent Neurology study found that cued recall was particularly effective in identifying 74 of 185 MCI patients whose memory impairment was related to underlying AD pathology using CSF A $\beta$ <sub>1-42</sub>/tau ratios, in contrast to a large number of MCI patients with non-specific pathology (Wagner et al., 2012). This was the first paper to provide biomarker evidence supporting longstanding clinical observations of the specificity of cued-recall memory impairment to AD.

We were interested in whether these findings could be applied earlier in the disease trajectory, that is, to preclinical AD stages. Stage 0 was defined as A $\beta$ -/ND-, Stage 1 as A $\beta$ + /ND-, and Stage 2 as A $\beta$ + /ND+ (Jack et al. 2012). An additional category of A $\beta$ -/ND+ clinically normal older adults were labeled as “Suspected Non-Alzheimer’s disease Pathology” (SNAP) by Jack and colleagues (Jack et al. 2012). We examined the relationship between preclinical stage and free and cued memory on a traditional verbal list learning task, the Selective Reminding Test (SRT) and the particularly challenging and experimental associative memory task, the MCT, in 260 clinically normal older adults participating in the Harvard Aging Brain Study. To determine whether these effects were specific to memory, we also examined the relationship between preclinical stage and global composites of memory, executive functioning, and processing speed.

## 2. MATERIALS AND METHODS

### 2.1 Sample characteristics

Our sample consisted of baseline data from the Harvard Aging Brain Study which was conducted at the Center for Alzheimer Research and Treatment at the Brigham and Women’s Hospital and Massachusetts General Hospital using protocols and informed consent procedures approved by the Partners Healthcare Human Research Committee. Subjects were deemed CN based on the following criteria: 1) a global Clinical Dementia Rating score of 0 (Morris, 1993) 2) scores above age and education-adjusted cutoffs on the Mini Mental State Exam (MMSE) (Folstein et al., 1975, Mungas et al., 1996, Crum et al., 1993) and on the 30-Minute Delayed Recall of the Logical Memory Story A (Wechsler, 1987, ADNI based cut-offs; <http://www.adni-info.org/>) and 3) a Geriatric Depression Scale (GDS) score of  $\leq$  11 (Yesavage et al., 1983). Review of medical history and physical and neurological examinations confirmed their status as clinically normal. None of the subjects had a history of alcoholism, drug abuse, head trauma, or current serious medical/psychiatric illness.

### 2.2 MRI data acquisition and analysis

Participants underwent MRI on a Siemens Trio-TIM 3 T scanner equipped with a 12-channel phased-array whole-head coil. High-resolution 3D T1-weighted multi-echo magnetization-prepared, rapid acquisition gradient echo anatomical images were collected with the following parameters: repetition time=2200ms; multi-echo echo times=1.54ms,

3.36ms, 5.18 ms, and 7 ms; flip angle 7°, 4x acceleration, 1.2×1.2×1.2 mm voxels. Region of interest (ROI) labeling was implemented using FreeSurfer v5.1. Hippocampus volume (HV) was collapsed across hemispheres and adjusted for estimated total intracranial volume (ICV): Adjusted HV (aHV) = raw HV—b (ICV —mean ICV), with b reflecting the regression coefficient when HV is regressed against ICV.

### 2.3 FDG-PET data acquisition and analysis

Fluorodeoxyglucose (FDG) PET imaging was completed at the MGH PET facility. Before injection, 10-minute transmission scans for attenuation correction were collected. 5.0–10.0 mCi was intravenously injected, and after a 45-min uptake period, FDG-PET images were acquired for 30-minutes in 3D acquisition mode.

FDG-PET data were realigned, summed, and normalized to a template using SPM8. FDG was extracted from a MetaROI reflecting AD vulnerable regions (lateral parietal, lateral inferior temporal and posterior cingulate cortex), and normalized by the mean from the top 50% of voxels from a pons/vermis reference region (Jagust et al., 2009).

### 2.4 Amyloid imaging acquisition and analysis

Amyloid burden was measured with N-methyl-[11C]-2-(4-methylaminophenyl)-6-hydroxybenzothiazole (Pittsburgh Compound B; PiB), which binds to fibrillar amyloid, and was prepared at Massachusetts General Hospital as described previously (Mathis et al., 2003; Klunk et al., 2004). Scans were completed at the MGH PET facility using a Siemens ECAT EXACT HR+ PET scanner. Before injection, 10-minute transmission scans for attenuation correction were collected. After injection of 8.5–15 mCi PIB, 60-minutes of dynamic data were acquired in 3D acquisition mode. Determination of A $\beta$  status is described elsewhere (Mormino et al., 2014a). In brief, PIB data were analyzed as standard uptake value ratios (SUVR), and a Gaussian mixture modeling approach was used to classify HABS CNs as A $\beta$ + or A $\beta$ - (cut-off value=1.20).

### 2.5 Classification into Neurodegeneration groups

Neurodegeneration (ND) status was determined based on aHV and MetaROI FDG (Jagust et al., 2009) as described by the Mayo Clinic (Jack et al., 2012). CNs were divided into ND+ and ND- groups based on cut-offs derived in a sample of ADNI AD patients, yielding cutoff values of 1.249 for the MetaROI FDG and 6723mm<sup>3</sup> for aHV. CNs were considered ND+ if they fell below the cut-off value for either ND marker (Jack et al., 2012).

### 2.6 Classification into preclinical stages based on joint A $\beta$ and ND status

Participants were classified into stages based on the presence/absence of A $\beta$  and ND whereby Stage 0 was defined as A $\beta$ -/ND-, Stage 1 as A $\beta$ + /ND-, and Stage 2 as A $\beta$ + /ND+ (Jack et al., 2012). An additional category of A $\beta$ - /ND+ CNs were labeled as “Suspected Non-Alzheimer’s disease Pathology” (SNAP) by Jack and colleagues (Jack et al., 2012). Although multiple groups have implemented these staging criteria (Knopman et al., 2012, Vos et al., 2013, Jack et al., 2012), differences among reports include derivation of A $\beta$  cut-offs, markers of ND (CSF tau versus imaging ND markers), and the use of subtle cognitive impairment as a categorization variable. Given our interest in determining the association

between preclinical stages and cognition, we excluded subtle cognitive impairment as a categorizing variable.

## 2.7 Neuropsychological tasks

Participants underwent an extensive initial cognitive assessment at baseline comprised of a total of 17 measures, some of which were standardized neuropsychological tasks and others experimental. Given a-priori hypotheses, there was a particular focus on memory tests. Because each cognitive task yields several variables, we employed a data reduction strategy to select variables to include in this report. To characterize cognition in relation to biomarker staging in the sample broadly, we used factor weightings from previously derived factor scores for executive functions, processing speed and memory (computed on 168 of the 260 HABS subjects included here) (Hedden et al., 2012). Each neuropsychological score was multiplied by the relevant factor weight from the prior report and these weighted scores were summed to compute the factor scores for processing speed, executive function, and memory. To ensure that the factor structure was not meaningfully altered by adding subjects to the prior report, we also conducted a comparison of the factor weightings in the completed baseline HABS sample and the prior sample from Hedden et al. (2012) and found convergence in the factor structure across samples.

The executive function factor score was composed of measures of verbal fluency (FAS/CAT; Benton et al., 1983, Monsch et al., 1992), letter-number sequencing from the Wechsler Adult Intelligence Scale-III (Wechsler, 1997), Digit Span Backward, Trail Making Test B minus A (Reitan, 1979), the Self-Ordered Pointing task, the Number Letter task, and a modified Flanker task (see Hedden et al., 2012 for full description and references). The processing speed factor score was composed of a reaction time measure from the Number Letter Task, Form A of the Trail Making Test, and the Digit-Symbol subtest of the WAIS-R (Wechsler, 1987). Factor scores were not computed for processing speed for 1 participant and for memory for 2 participants due to missing data.

The memory factor score included the name and occupation recall on the Face Name Associative Memory Exam (Rentz et al., 2010), delayed free recall on the 6-Trial Selective Reminding Test (Masur et al., 1990), and delayed free recall of List 2 on the Memory Capacity Test (H. Buschke, personal communication; Rentz et al., 2010). Because we were particularly interested in examining component parts of memory, we examined delayed recall on the MCT and SRT, both of which are included in the memory factor score but both of which are also designed with free and facilitated recall trials. Additionally, both the MCT and SRT are challenging and thus potentially more sensitive to subtle changes in memory in a non-clinical population. However, they are challenging in different ways; while the MCT involves learning and remembering significantly more items (32 compared with 12), learning is facilitated on the MCT with cues whereas learning on the SRT is made more challenging by only providing forgotten words trial by trial. In this way, they may be differentially challenging, but allow for a comparison between associative memory (MCT) and list-based learning (SRT).

The MCT is an associative memory measure, which pairs a written word with a semantic category to enhance encoding and retrieval. A total of 32 words are learned, in two lists of

16 items (see Figure 1). Both lists contain words from the same 16 semantic categories (e.g., country, color) but are unique (e.g. Spain is in list 1 and Italy is in list 2) and learned separately. In the study phase, the words from List 1 are presented 4 per slide; the examiner names a category (e.g., country) and the subject is required to name the corresponding word from the 4 available (e.g., Spain). After identifying all 16 items according to the semantic cue, the examiner verbally presents the cue alone and asks the participant to recall all words. This procedure is repeated for List 2. Following this stage, the participant is provided with the semantic cue once again, but is now asked to produce the items from both lists (Immediate Cued Recall). This is followed by Immediate Free Recall in which the participant produces remembered items from both lists without the benefit of a cue. Following a 30-minute delay, the participant is again asked to freely produce remembered items from both lists, resulting in a Delayed Free Recall score of both lists (/32) followed by a Delayed Cued Recall (/32).

The SRT is a measure of verbal list-learning and memory. Participants are asked to remember 12 orally-presented words over 6 trials but, following the first learning trial, they are only “selectively reminded of” those words omitted on the immediately preceding trial. Total learning over the 6 trials results in a maximum score of 72. Following a 10-minute delay, participants are asked to recall the words freely resulting in a free recall score (maximum of 12) and a recognition (referred to here for ease of interpretation as cued recall) score (maximum of 12) where subjects are asked to select the correct word from a 4-item multiple choice paradigm. At least 1 of the foils is either semantically or phonemically related to the target (for example, bell is one of the options where bowl is the target and sunrise is one of the options where dawn is the target).

## 2.8 Statistical Analyses

Statistical analyses were completed using SPSS v22. Differences in demographic variables and cognitive screening measures (MMSE, Logical Memory II) across the 4 biomarker-defined stages were examined with univariate linear regression equations and subsequent pairwise t-tests for continuous variables (age, years of education) and chi-squared tests for dichotomous variables (sex). For ease of comparison, descriptive statistics for cognitive variables are presented as z-scores, derived from whole-sample means and standard deviations for each corresponding variable. The analyses described below used a standard set of covariates consisting of sex, age, and years of education.

Separate univariate linear regression equations for each of the factors scores (executive functioning, processing speed, and memory) were computed, adjusting for the standard set of covariates. Post-hoc pairwise t-tests were then performed to make comparisons between stages. Separate univariate linear regression equations were then computed using the standard covariates for selected measures of delayed recall in the memory factor score (MCT free recall, MCT cued recall, SRT free recall, SRT cued recall). Because some reports have shown that initial declines in episodic memory occur at the stage of learning (Grober & Kawas, 1997, Bilgel et al. 2014), we completed secondary analyses examining total learning on the SRT and immediate free and cued recall on the MCT. For non-normally distributed cognitive variables (MCT cued recall and SRT cued recall) a log transformation was applied

and used in subsequent analyses. We decided to use the SRT cued score, although transforming it did not create a normal distribution, given the relative robustness of ANOVA to deviations in normality (Maxwell & Delaney, 2004). For all analyses, data were screened for outliers, with no cases exhibiting standardized residuals exceeding  $\pm 3.0$  standard deviations. Standardized residuals for the stages and for the overall model were assessed for normality using Shapiro-Wilk's test ( $p > .05$ ). Homoscedasticity and homogeneity of variances were assessed by visual inspection of scatterplots and Levene's test ( $p > .05$ ), respectively.

### 3. RESULTS

A total of 260 subjects had all biomarker data available to be classified into stages. The number of subjects assigned to each biomarker stage and demographic variables are shown in Table 1. There were group differences in age [ $F(3,1)=12.172$ ,  $p<0.0001$ ] whereby Stage 2 participants were older compared with Stage 0 ( $p<0.0001$ ) and Stage 1 ( $p=0.005$ ). Similarly SNAP participants were older compared with participants in Stage 0 ( $p<0.0001$ ). There were no differences in education level [ $F(3,1)=0.663$ ,  $p=0.575$ ], MMSE [ $F(3,1)=2.056$ ,  $p=0.107$ ], or Logical Memory Delayed Recall scores [ $F(3,1)=0.347$ ,  $p=0.791$ ] amongst the 4 groups. SNAP was comprised of more males compared with Stage 0 [ $\chi^2(1,191)=4.476$ ,  $p=0.034$ ].

#### 3.1 Memory

#### 3.2 Overall Group Effects

There was a trend toward a significant group effect on overall memory factor score performance while including standard covariates [ $F(3, 252)=2.427$ ,  $p=0.066$ ,  $R^2 = 0.146$ ] (see Figure 2). Although this overall model only reaches a trend level of significance, we pursued further pairwise comparisons given this trend and a-priori hypotheses regarding the relationship between these biomarker stages and memory. When examining the individual tests, we found a relationship between biomarker staging and MCT free recall [ $F(3,254) = 3.367$ ,  $p = 0.019$ ,  $R^2 = 0.098$ ] and biomarker staging and MCT cued recall [ $F(3,254) = 2.732$ ,  $p = 0.044$ ,  $R^2 = 0.074$ ] (see Figure 3). There was no difference amongst stages on MCT immediate recall [ $F(3,254) = 1.729$ ,  $p = 0.161$ ,  $R^2 = 0.106$ ] nor on MCT immediate cued recall [ $F(3,254) = 1.681$ ,  $p = 0.172$ ,  $R^2 = 0.053$ ]. Biomarker group effects were not observed for SRT free recall [ $F(3,253) = 0.965$ ,  $p = 0.410$ ,  $R^2 = 0.046$ ] nor SRT cued [ $F(3,253) = 0.126$ ,  $p = 0.945$ ,  $R^2 = -0.013$ ] (see Figure 3). In addition, there were no overall group effects on SRT total learning [ $F(3,252) = 0.996$ ,  $p = 0.395$ ,  $R^2 = 0.081$ ].

#### 3.3 Stage 0 vs. 1

Subjects with amyloidosis alone ( $A\beta+/ND-$ ) performed worse on MCT free recall compared with those without biomarker evidence of pathology ( $p=0.030$ ) ( $A\beta-/ND-$ ), see Figure 2. However, there were no significant group differences between Stages 0 and 1 on the memory factor score (0.345), MCT cued recall ( $p=0.880$ ), nor SRT free recall ( $p=0.970$ ) and cued recall ( $p=0.966$ ).

### 3.4 Stage 0 vs. 2

On the overall memory factor score, participants classified as positive for both A $\beta$  and ND performed worse compared with those in Stage 0 (A $\beta$ -/ND-,  $p=0.008$ ). When examining individual tests, Stage 2 scored lower compared with Stage 0 for both MCT free recall ( $p=0.009$ ) and MCT cued recall ( $p=0.006$ ). However, on the SRT, there were no Stage 0 vs. 2 differences on free recall ( $p=0.129$ ) or cued recall ( $p=0.578$ ) though there appears to be the beginning of a trend for SRT free recall (see Figure 3).

### 3.5 Stage 0 vs. SNAP

For MCT free recall, there was a trend toward the SNAP group performing worse compared with Stage 0 ( $p=0.062$ ) but no differences in MCT cued recall ( $p=0.152$ ). There were otherwise no significant group differences between Stages 0 and SNAP on the memory factor score ( $p=0.245$ ), MCT cued recall ( $p=0.238$ ), or SRT free recall ( $p=0.862$ ) and cued recall ( $p=0.745$ ).

### 3.6 Stage 1 vs. Stage 2

For MCT cued recall, Stage 2 performed worse compared with both Stage 0 and Stage 1 ( $p=0.036$ ). There were no differences between these stages for MCT free recall (0.741), the memory factor score ( $p=0.174$ ), or SRT free recall ( $p=0.235$ ) or cued recall ( $p=0.631$ ).

### 3.7 Executive Functioning

There were no differences in executive functioning in relation to biomarker stage using the standard covariates [ $F(3, 254)=0.652, p=0.582$ ] (see Figure 2).

### 3.8 Processing Speed

There were no group differences in processing speed in relation to biomarker stage using the standard covariates [ $F(3, 253)=0.361, p=0.781$ ] (see Figure 2).

## 4. DISCUSSION

Our findings indicate that there are subtle, yet measurable memory decrements in normal older adults with biomarker evidence of preclinical Alzheimer's disease (AD). These biomarker-cognition relationships are more likely to be detected in otherwise clinically normal individuals using particularly challenging and specific measures, such as the Memory Capacity Test (MCT). The combination of amyloidosis and neurodegeneration (ND) in Stage 2 is associated with more reliably detectable and more advanced memory decrements (a decline in both free and cued recall). Older adults with amyloidosis alone (Stage 1) exhibit reductions in MCT free recall only, even when controlling for their age, sex, and educational achievement. Interestingly, ND alone (SNAP) is associated with a trend towards lower free recall compared with Stage 0, but ND alone does not appear to be sufficient to produce cued memory decrements. Taken together, these findings lend some support for the diagnostic specificity of impaired cued recall to underlying AD pathology. However, it also suggests a temporal ordering of memory decline whereby decrements in free recall are a leading indicator in preclinical AD, followed by decrements in cued recall. It also suggests that challenging associative memory tasks may be more useful at detecting



cross-sectional performance differences (MCT) compared with more traditional verbal learning tasks (SRT). We did not find a relationship between biomarker group staging and global composites of executive functioning or processing speed.

There is longstanding clinical evidence, which suggests that memory failure that persists despite cueing represents medial temporal storage dysfunction secondary to AD pathology (Moss et al., 1986). Associative memory tests such as the MCT were designed to isolate medial temporal memory storage capacities by facilitating encoding and retrieval with semantic cueing, thus limiting the confounding effects of attention and processing speed known to decline with age (Grober et al., 2000). In this way, associative memory paradigms, in contrast with more traditional verbal learning tasks, may be less susceptible to the moderating effect of compensatory strategies, such as clustering and organization, on performance. Individuals with a medial-temporal amnesia will perform poorly on both free and cued recall whereas individuals whose memory difficulties are related to more subcortico-frontal processes tend to perform poorly on free recall but normally on cued recall (Grober et al. 2010). These hypotheses are consistent with recent empirical work in the clinical literature showing that cued memory differentiates normal aging from amnesic MCI (Bennett et al., 2006, Algarabel et al., 2012, Wolk et al., 2013), differentiates progressive versus stable memory impairment (Dierckx et al., 2007, Sarazin et al., 2007) and most interestingly, differentiates memory impairment associated with biomarker evidence of prodromal AD versus non-AD related Mild Cognitive Impairment (Wagner et al., 2012).

In line with these findings, we observed an association between decrements in cued recall and the presence of advancing pathology that included both amyloidosis and ND (Stage 2). While some groups have proposed refining the definition of episodic memory decline related to AD from ‘decrements in free recall’ to ‘decrements in free recall unimproved with cueing’ (Dubois et al., 2007), this definition may not completely capture the earliest changes associated with preclinical AD. More specifically, only reduced free recall was associated with amyloidosis alone (Stage 1). Decrements in cued recall were not observed in Stage 1 or SNAP, suggesting that amyloidosis or ND in isolation may be insufficient to impact cued recall. This is consistent with a growing body of work suggesting that the combination of amyloidosis and ND results in greater cognitive decrements and greater risk of clinical progression (Knopman et al., 2012, Vos et al., 2013, Mormino et al., 2014a, Mormino et al., 2014b). It also provides evidence for the temporal ordering of memory decline in preclinical AD whereby a decline in free recall may be a leading indicator and a decline in cued recall represents further progression.

Interestingly, we did not find stage-based performance differences on the SRT. The MCT may be more sensitive cross-sectionally because of its more challenging design (Rentz et al. 2013, Pike et al. 2011) in the number of items to be learned and the length of the delayed recall. In addition, it may be that individuals without significant underlying pathology benefited more from the structured learning on the MCT, thus maximizes the differences between biomarker groups. The selective reminding procedure for learning items on the SRT makes it challenging (e.g., not all items are repeated at each learning trial), but there are significantly fewer items to encode on the SRT (12 total) compared with the 32 required on the MCT. In addition, the associative paradigm and the semantic cueing of the MCT may

be more specifically targeting medial-temporal processes compared with more general verbal memory paradigms such as the SRT. The cued recall paradigm on the MCT is also less susceptible to ceiling effects and subsequent restricted variance which is common to tasks of memory familiarity. In our study, significant ceiling effects were observed on the SRT (55% of subjects obtained perfect scores on the multiple choice/recognition portion). In contrast, only 7% of subjects scored at ceiling level on cued recall on the MCT. As such, the MCT provides increased performance variance which is especially important in assessing CN older adults, where memory changes are subtle. We additionally examined total learning on both the MCT and SRT because some reports have shown that primary declines in episodic memory occur at the stage of learning (Grober & Kawas, 1997, Bilgel et al. 2014) while secondary declines occur at delayed recall. However, the lack of group differences between biomarker stages on initial learning on both the SRT and MCT suggest that delayed recall on the MCT is more sensitive compared with initial learning. Because significant differences were not found on either initial learning or delayed recall of the SRT, it is unclear if delayed recall is more sensitive compared with initial learning to preclinical AD across all memory measures.

It will be important to further tease apart the relationships between the component processes of memory and biomarkers—for example, whether semantic cueing is more sensitive compared with multiple choice or yes/no paradigms, where the earliest memory breakdowns occur (i.e. at stage of encoding vs. retrieval) longitudinally, and whether second list learning may be particularly sensitive. Doing so may be a vital means of identifying the metrics best able to either detect and/or track subtle memory changes prior to expensive PET and LP biomarker collection in secondary AD prevention trials (see Friedrich, 2014 for overview of AD prevention trials).

While we did not find a relationship between biomarker group staging and global composites of executive functioning or processing speed, other groups have shown correlations between these cognitive domains and lower FDG metabolism in AD-ROIs and smaller hippocampal volumes in clinically normal adults; however, these studies examined biomarkers as continuous variables (Mielke et al., 2014). Given that the earliest cognitive changes associated with AD occur within the memory domain, it is unsurprising that executive functions and processing speed are not associated with amyloid pathologies in clinically normal older adults. Further understanding the differences between memory decline in Stage 2 versus SNAP, especially with longitudinal data, will be vital to understanding not only the earliest AD-related changes in memory but whether declines in memory attributed to “normal aging” are in fact benign.

#### 4.1 Limitations and Future Directions

The subtle cognitive decline described in the preclinical criteria as Stage 3 may be detectable at Stage 1 on more challenging and specific tests of memory (here, the MCT). However, changes in memory are more reliably detected in individuals in Stage 2. Furthering our understanding of the cross sectional cognitive profiles associated with preclinical stages may potentially inform decisions about the most appropriate cognitive inclusion and outcome measures to identify the earliest subtle but worrisome memory

changes. A larger and more heterogeneous validation sample is required to establish appropriate sensitivity, specificity and cut-off scores on the MCT in relation to AD biomarkers. In addition, it will be important in future work to determine if free vs. cued memory decline has differential sensitivity to longitudinal memory decline and clinical progression.

The Harvard Aging Brain study is uniquely designed in that it involves the administration of multiple memory paradigms, some of which are traditional and others experimental. Comparing memory tests against each other highlights the importance of test selection in clinical assessment and trials, especially as we move earlier in the AD trajectory, where detecting subtle yet worrisome memory changes is a challenge.

## Acknowledgments

### FUNDING:

NIA/NIH: P01AG036694 (RAS), P50 AG005134 (TH), 5T32AG023480-08 (KVP), K01 AG040197 (TH), F32AG044054 (EM), K23AG044431 (REA), U01 AG032438 (RAS), U01 AG024904 (RAS), R01 AG037497 (RAS, KAJ), R01 AG034556 (RAS, KAJ), K24 AG035007, P50 AG0051341 (KAJ, RAS), U19 AG010483(KAJ, RAS), R01 AG027435 (KAJ, RAS) and P01 AG036694 (RAS, KAJ). The Alzheimer's Association: COG-13-282201 (DMR), NIRG-12-243012 (REA). The Charles King Trust (KVP)

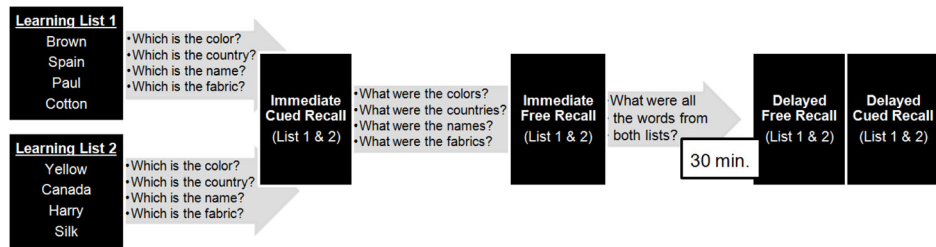
The authors wish to thank the invaluable Harvard Aging Brain study participants who make this work possible. We also thank our dedicated research assistants including: Sarah Aghjayan, Margaret Chute, Alex Dagley, Maria Dekhtyar, Sehily Jaimes, Molly Lapoint, Tamy-Fee Meneide, Catherine Munro, and Sarah Wigman and co-investigators Jasmeer Chhatwal, Nancy Donovan, Willem Huijbers, Gad Marshall, Donald McLaren, Aaron Schultz, Patrizia Vannini and Yakeel Quiroz. In addition, a special thanks to Dr. Herman Buschke and the Albert Einstein College of Medicine of Yeshiva University for allowing us to use the Memory Capacity Test.

## References

- Algarabel S, Fuentes M, Escudero J, Pitarque A, Peset V, Mazón JF, Meléndez JC. Recognition memory deficits in mild cognitive impairment. *Aging, Neuropsychology, and Cognition*. 2012; 19(5):608–619.
- Bennett DA, Schneider JA, Arvanitakis Z, Kelly JF, Aggarwal NT, Shah RC, Wilson RS. Neuropathology of older persons without cognitive impairment from two community-based studies. *Neurology*. 2006; 66(12):1837–1844. [PubMed: 16801647]
- Benton, AL., editor. *Contributions to neuropsychological assessment: A clinical manual*. Oxford University Press; 1994.
- Bilgel M, An Y, Lang A, Prince J, Ferrucci L, Jernigan B, Resnick SM. Trajectories of Alzheimer disease-related cognitive measures in a longitudinal sample. *Alzheimer's & Dementia*. 2014; 10(6): 735–742.
- Crum RM, Anthony JC, Bassett SS, Folstein MF. Population-based norms for the Mini-Mental State Examination by age and educational level. *Jama*. 1993; 269(18):2386–2391. [PubMed: 8479064]
- Dierckx E, Engelborghs S, De Raedt R, De Deyn PP, Ponjaert-Kristoffersen I. Differentiation between mild cognitive impairment, Alzheimer's disease and depression by means of cued recall. *Psychological medicine*. 2007; 37(05):747–755. [PubMed: 17164030]
- Dubois B, Feldman HH, Jacova C, Cummings JL, DeKosky ST, Barberger-Gateau P, Scheltens P. Revising the definition of Alzheimer's disease: a new lexicon. *The Lancet Neurology*. 2010; 9(11): 1118–1127. [PubMed: 20934914]
- Dubois B, Feldman HH, Jacova C, DeKosky ST, Barberger-Gateau P, Cummings J, Scheltens P. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *The Lancet Neurology*. 2007; 6(8):734–746. [PubMed: 17616482]

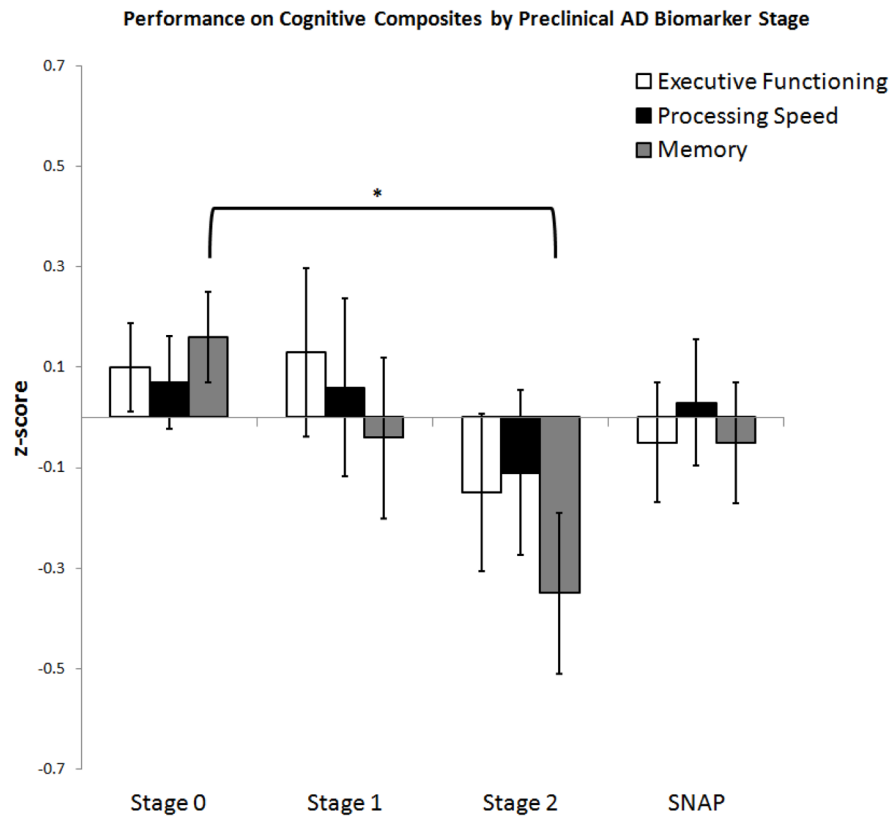
- Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*. 1975; 12(3):189–198. [PubMed: 1202204]
- Friedrich MJ. Researchers test strategies to prevent Alzheimer disease. *JAMA*. 2014; 311(16):1596–1598. [PubMed: 24718971]
- Grober E, Kawas C. Learning and retention in preclinical and early Alzheimer's disease. *Psychology and aging*. 1997; 12(1):183. [PubMed: 9100279]
- Grober E, Sanders AE, Hall C, Lipton RB. Free and cued selective reminding identifies very mild dementia in primary care. *Alzheimer disease and associated disorders*. 2010; 24(3):284. [PubMed: 20683186]
- Grober E, Lipton RB, Hall C, et al. Memory impairment on free and cued selective reminding predicts dementia. *Neurology*. 2000; 54:827–832. [PubMed: 10690971]
- Hedden T, Mormino EC, Amariglio RE, Younger AP, Schultz AP, Becker JA, Rentz DM. Cognitive profile of amyloid burden and white matter hyperintensities in cognitively normal older adults. *The Journal of Neuroscience*. 2012; 32(46):16233–16242. [PubMed: 23152607]
- Jack CR, Knopman DS, Weigand SD, Wiste HJ, Vemuri P, Lowe V, Petersen RC. An operational approach to National Institute on Aging–Alzheimer's Association criteria for preclinical Alzheimer disease. *Annals of neurology*. 2012; 71(6):765–775. [PubMed: 22488240]
- Jagust WJ, Landau SM, Shaw LM, Trojanowski JQ, Koeppe RA, Reiman EM, Mathis CA. Relationships between biomarkers in aging and dementia. *Neurology*. 2009; 73(15):1193–1199. [PubMed: 19822868]
- Klunk WE, Engler H, Nordberg A, Wang Y, Blomqvist G, Holt DP, Långström B. Imaging brain amyloid in Alzheimer's disease with Pittsburgh Compound- B. *Annals of neurology*. 2004; 55(3):306–319. [PubMed: 14991808]
- Knopman DS, Jack CR, Wiste HJ, Weigand SD, Vemuri P, Lowe V, Petersen RC. Short-term clinical outcomes for stages of NIA-AA preclinical Alzheimer disease. *Neurology*. 2012; 78(20):1576–1582. [PubMed: 22551733]
- Mathis CA, Wang Y, Holt DP, Huang GF, Debnath ML, Klunk WE. Synthesis and evaluation of 11C-labeled 6-substituted 2-arylbenzothiazoles as amyloid imaging agents. *Journal of medicinal chemistry*. 2003; 46(13):2740–2754. [PubMed: 12801237]
- Masur DM, Fuld PA, Blau AD, Crystal H, Aronson MK. Predicting development of dementia in the elderly with the Selective Reminding Test. *Journal of Clinical and Experimental Neuropsychology*. 1990; 12(4):529–538. [PubMed: 2211975]
- Maxwell, SE.; Delaney, HD. *Designing experiments and analyzing data: A model comparison perspective*. Vol. 1. Psychology Press; 2004.
- Mielke MM, Weigand SD, Wiste HJ, Vemuri P, Machulda MM, Knopman DS, Petersen RC. Independent comparison of CogState computerized testing and a standard cognitive battery with neuroimaging. *Alzheimer's & Dementia*. 2014; 10(6):779–789.
- Morris, John C. *The Clinical Dementia Rating (CDR): current version and scoring rules*. *Neurology*. 1993
- Monsch AU, Bondi MW, Butters N, Salmon DP, Katzman R, Thal LJ. Comparisons of verbal fluency tasks in the detection of dementia of the Alzheimer type. *Archives of Neurology*. 1992; 49(12):1253–1258. [PubMed: 1449404]
- Moss MB, Albert MS, Butters N, Payne M. Differential patterns of memory loss among patients with Alzheimer's disease, Huntington's disease, and alcoholic Korsakoff's syndrome. *Archives of Neurology*. 1986; 43(3):239–246. [PubMed: 2936323]
- Mormino EC, Betensky RA, Hedden T, Schultz AP, Amariglio RE, Rentz DM, Sperling RA. Synergistic Effect of  $\beta$ -Amyloid and Neurodegeneration on Cognitive Decline in Clinically Normal Individuals. *JAMA neurology*. 2014; 71(11):1379–1385. [PubMed: 25222039]
- Mormino EC. The Relevance of Beta-Amyloid on Markers of Alzheimer's Disease in Clinically Normal Individuals and Factors That Influence These Associations. *Neuropsychology review*. 2014; 24(3):300–312. [PubMed: 25108368]

- Mungas D, Marshall SC, Weldon M, Haan M, Reed BR. Age and education correction of Mini-Mental State Examination for English-and Spanish-speaking elderly. *Neurology*. 1996; 46(3):700–706. [PubMed: 8618670]
- Pike KE, Ellis KA, Villemagne VL, Good N, Chételat G, Ames D, Rowe CC. Cognition and beta-amyloid in preclinical Alzheimer's disease: data from the AIBL study. *Neuropsychologia*. 2011; 49(9):2384–2390. [PubMed: 21529702]
- Reitan, RM. Manual for Administration of Neuropsychological Test Batteries for Adults and Children. Tuscon, AZ: Reitan Neuropsychology Laboratory; 1979.
- Rentz DM, Locascio JJ, Becker JA, Moran EK, Eng E, Buckner RL, Johnson KA. Cognition, reserve, and amyloid deposition in normal aging. *Annals of neurology*. 2010; 67(3):353–364. [PubMed: 20373347]
- Rentz DM, Parra Rodriguez MA, Amariglio R, Stern Y, Sperling R, Ferris S. Promising developments in neuropsychological approaches for the detection of preclinical Alzheimer's disease: a selective review. *Alzheimers Res Ther*. 2013; 5:58. [PubMed: 24257331]
- Sarazin M, Berr C, De Rotrou J, Fabrigoule C, Pasquier F, Legrain S, Dubois B. Amnesic syndrome of the medial temporal type identifies prodromal AD A longitudinal study. *Neurology*. 2007; 69(19):1859–1867. [PubMed: 17984454]
- Sperling RA, Jack CR, Aisen PS. Testing the right target and right drug at the right stage. *Science translational medicine*. 2011b; 3(111):111cm33–111cm33.
- Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Phelps CH. Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*. 2011a; 7(3):280–292.
- Vos SJ, Xiong C, Visser PJ, et al. Preclinical Alzheimer's disease and its outcome: a longitudinal cohort study. *Lancet Neurol*. 2013; 12:957–965. [PubMed: 24012374]
- Wagner M, Wolf S, Reischies FM, Daerr M, Wolfsgruber S, Jessen F, Wiltfang J. Biomarker validation of a cued recall memory deficit in prodromal Alzheimer disease. *Neurology*. 2012; 78(6):379–386. [PubMed: 22238414]
- Wechsler, D. Wechsler Memory Scale-Revised. San Antonio: The Psychological Corporation; 1987.
- Wechsler, D. WAIS-III, Wechsler adult intelligence scale administration and scoring manual. 3. New York: The Psychological Corporation; 1997.
- Wolk DA, Mancuso L, Kliot D, Arnold SE, Dickerson BC. Familiarity-based memory as an early cognitive marker of preclinical and prodromal AD. *Neuropsychologia*. 2013; 51(6):1094–1102. [PubMed: 23474075]
- Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO. Development and validation of a geriatric depression screening scale: a preliminary report. *Journal of psychiatric research*. 1983; 17(1):37–49. [PubMed: 7183759]

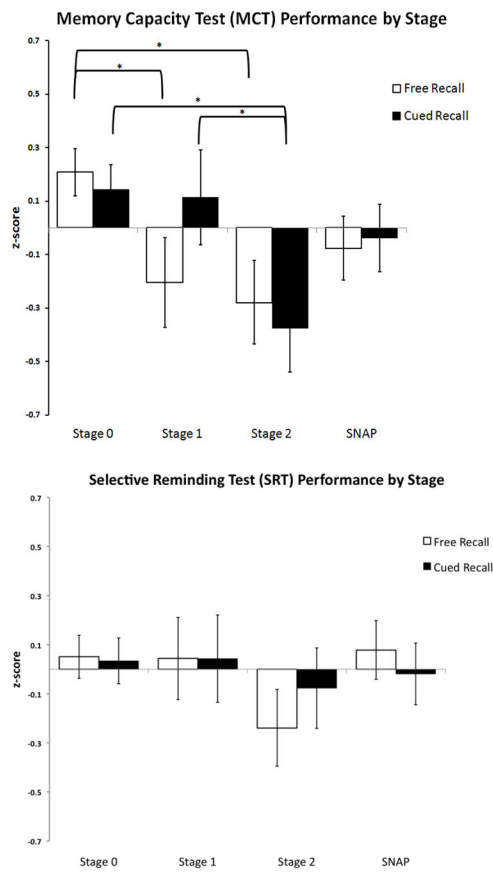


**Figure 1.**

The Memory Capacity Test. In List Learning 1, subjects view 4 words per page. The examiner provides a semantic cue (e.g., which item is a country?) for each item and the subject selects the correct corresponding word (e.g. Spain). This is repeated for 3 more cards, for a total of 16 words comprising List1. This procedure is repeated for a 2<sup>nd</sup> list of 16 new words but with equivalent semantic categories (List Learning 2). A semantic cue is provided and subjects are asked to produce both words from List 1 and 2 (Immediate Cued Recall). This is followed by Immediate Free Recall in which the subject produces items from both lists without a cue. The recall portion is repeated following a 30-minute delay with Delayed Free Recall (/32) followed by Delayed Cued Recall (/32).



**Figure 2.** Performance on Cognitive Composites by Preclinical AD Biomarker Stage. Estimated marginal means (z-scores) and standard errors by Stage controlled for age, sex, and education. \*Brackets show the significant group differences ( $p < 0.05$ ) between Stage 0 and Stage 2 on memory performance.



**Figure 3.** Memory Performance on Individual Tests (MCT, SRT) with Free and Cued Paradigms by Biomarker Stage. Estimated marginal means (z-scores) and standard errors (controlling for age, sex, and education) for MCT Delayed Recall (top) and SRT Delayed Recall by biomarker stage. \*Brackets represent significant ( $p < 0.05$ ) pairwise comparisons.



**Table 1**

Subject Demographics and Global Cognition. Medians and Interquartile ranges, unless otherwise stated, are listed for demographic variables by Stage.

	<b>Stage 0</b> n=126 A $\beta$ -/ND-	<b>Stage 1</b> n=32 A $\beta$ + /ND-	<b>Stage 2</b> n=38 A $\beta$ + /ND+	<b>SNAP</b> n=64 A $\beta$ - /ND+
Age*	71 (67–76)	72 (69–78)	77 (74–83)	76 (70–81)
Sex % male*	34.9%	40.6%	36.8%	50.8%
Education (years)	16 (13–18)	16 (14–18)	16 (14–18)	16 (12–18)
MMSE <sup>^</sup>	29 (29–30)	29 (28–30)	29 (28–29)	29 (28–30)
Logical Memory Delayed Recall	14 (12–16)	15 (12–17)	14 (12–16)	14 (12–16)

\* group differences p<0.05,

<sup>^</sup> trend level significance